

2A

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7:		(11) International Publication Number: WO 00/40695	
C12N		(43) International Publication Date: 13 July 2000 (13.07.00)	
(21) International Application Number: PCT/US00/0466		(74) Agents: STEWART, Raymond, C. et al.; Birch, Stewart, Kolsch & Birch, LLP, P.O. Box 747, Falls Church, VA 22040-0747 (US).	
(22) International Filing Date: 7 January 2000 (07.01.00)			
(30) Priority Data: 60/115,293 8 January 1999 (08.01.99) US			
(71) Applicant (for all designated States except US): CERES, INC. (US/US); 3007 Malibu Canyon Road, Malibu, CA 90265 (US).			
(72) Inventors; and (75) Inventors/Applicants (for US only): ALEXANDROV, Nikolai [RU/US]; 1404 Oak Trail Street, Thousand Oaks, CA 91320 (US); BROVER, Vyacheslav [RU/US]; 5916 N. Las Virgenes Road #390, Calabasas, CA 91320 (US); CHEN, Xianfeng [CN/US]; 12333 Wild Turkey Court, #B, Creve Coeur, MO 63141 (US); SUBRAMANIAN, Gopalakrishnan [IN/US]; 4205 Peach Slope Road, Moorpark, CA 93021 (US); TROUKHAN, Maxim R. [RU/US]; 1675 Amberwood Drive #2, South Pasadena, CA 91030 (US); ZHENG, Lian-sheng [CN/US]; 19212 Circle Gate Drive #201, Germantown, MD 20874 (US).			
Published Without international search report and to be republished upon receipt of that report.			
(54) Title: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES ENCODED THEREBY			
(57) Abstract The present invention provides DNA molecules that constitute portions of the genome of a plant, and polypeptides encoded thereby. The DNA molecules are useful for expressing a gene product, either as a promoter or as a structural gene or as an UTR or as a 3' terminal sequence and are also useful in controlling expression of the target gene or as tools for genetic mapping or identification of a particular individual plant or for clustering of a group of plants with a common trait.			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AN	Andorra	FR	France	LV	Latvia	SN	Senegal
AT	Austria	GB	United Kingdom	LU	Luxembourg	SZ	Switzerland
AU	Australia	GA	Gabon	MD	Moldova	TD	Togo
AZ	Azerbaijan	GE	Georgia	MG	Madagascar	TG	Togo
BA	Bosnia and Herzegovina	GH	Ghana	MK	Macedonia	TM	Turkmenistan
BB	Barbados	GN	Guinea	ML	Mali	TT	Trinidad and Tobago
BE	Belgium	GR	Greece	MN	Mongolia	UA	Ukraine
BF	Burkina Faso	HR	Croatia	MR	Mauritania	UG	Uganda
BG	Bulgaria	IE	Ireland	MW	Malawi	US	United States of America
BJ	Benin	IL	Israel	MX	Mexico	UZ	Uzbekistan
BR	Brazil	IS	Iceland	NE	Niger	VN	Viet Nam
BT	Bhutan	IT	Italy	NL	Netherlands	YU	Yugoslavia
CA	Canada	JP	Japan	NO	Norway	ZW	Zimbabwe
CC	Cocos (Keeling) Islands	KG	Kyrgyzstan				
CD	Congo	KP	Democratic People's Republic of Korea				
CF	Cote d'Ivoire	KR	Republic of Korea				
CG	Congo	KZ	Kazakhstan				
CH	Switzerland	LC	St. Lucia				
CN	China	LI	Liechtenstein				
CU	Cuba	LU	Luxembourg				
CZ	Czech Republic	LV	Latvia				
DE	Germany	MA	Morocco				
DK	Denmark	MC	Monaco				
EE	Estonia	ME	Montenegro				

**SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING  
POLYPEPTIDES ENCODED THEREBY**

**FIELD OF THE INVENTION**

5 The present invention relates to isolated polynucleotides that encode all, or a portion of, a gene that is expressed and the corresponding polypeptide. The present invention also relates to isolated polynucleotides that encode regulatory regions of genes.

**DESCRIPTION OF THE RELATED ART**

10 Efforts to map and sequence the genome of a number of organisms are in progress; a few complete genome sequences, for example those of *E. coli* and *Saccharomyces cerevisiae* are known (Blattner et al., Science 277:1453 (1997); Goffeau et al., Science 274:546 (1996)). The complete genome of a multicellular organism, *C. elegans*, has also been sequenced (See, the *C. elegans* Sequencing Consortium, Science 282:2012 (1998)). To date, no complete genome of a plant has been sequenced, nor has a complete cDNA complement of any plant been sequenced.

**SUMMARY OF THE INVENTION**

20 The present invention comprises polynucleotides, such as complete cDNA sequences and/or sequences of genomic DNA encompassing complete genes, portions of genes, and/or intergenic regions, hereinafter collectively referred to as "Sequence-Determined DNA Fragments" (SDFs), from plants, particularly corn and *Arabidopsis thaliana* and polypeptides derived therefrom. In some instances, the SDFs span the entirety of a protein-coding segment. In some instances, the entirety of an mRNA is represented. Other objects of the invention are the control sequences, such as but not limited to promoters, that are also represented by SDFs of the invention. Complements of any sequence of the invention are also considered part of the invention.

Other objects of the invention are polynucleotides comprising exon sequences, polynucleotides comprising intron sequences, polynucleotides comprising introns together with exons, intron/exon junction sequences, 5' untranslated sequences, and 3' untranslated sequences of the SDFs of the present invention. Polynucleotides representing the joiner of any exons described herein, in any arrangement, for example, to produce a sequence encoding any desirable amino acid sequence are within the scope of the invention.

10 The present invention also resides in probes useful for isolating and identifying nucleic acids that hybridize to an SDF of the invention. The probes are typically of a length of 12 to 2000 nucleotides long; more typically, 15 to 200 nucleotides long; even more typically, 18 to 100 nucleotides long.

Yet another object of the invention is a method of isolating and/or identifying nucleic acids using the following steps:

- (a) contacting a probe of the instant invention with a polynucleotide sample under conditions that permit hybridization and formation of a polynucleotide duplex; and
- (b) detecting and/or isolating the duplex of step (a).

The conditions for hybridization can be from low to moderate to high stringency conditions. The sample can include a polynucleotide having a sequence unique in a plant genome. Probes and methods of the invention are useful, for example, without limitation, for mapping of genetic traits and/or for positional cloning of a desired portion of genomic DNA.

Probes and methods of the invention can also be used for detecting alternatively spliced messages within a species. Probes and methods of the invention can further be used to detect or isolate related genes in other plant species using genomic DNA (gDNA) and/or cDNA libraries. In some instances, especially when longer probes and low to moderate stringency



3

hybridization conditions are used, the probe will hybridize to a plurality of cDNA and/or gDNA sequences of a plant. This approach is useful for isolating representatives of gene families which are identifiable by possession of a common functional domain in the gene product or which have common cis-acting regulatory sequences. This approach is also useful for identifying orthologous genes from other organisms, which can be more or less related to corn, Arabidopsis, or another plant.

The present invention also resides in constructs for modulating the expression of the genes comprised of all or a portion of an SDF. The constructs comprise all or a portion of the expressed SDF, or of a complementary sequence. Examples of constructs include ribozymes comprising RNA encoded by an SDF or by a sequence complementary thereto, antisense constructs, constructs comprising coding regions or parts thereof, constructs comprising promoters, introns, untranslated regions, etc. When inserted into a host cell the construct is, preferably, functionally integrated with or operatively linked to a heterologous polynucleotide. For instance, a coding region from an SDF might be operably linked to a promoter that is functional in a plant.

The present invention also resides in host cells, including bacterial or yeast cells or plant cells, and transgenic plants that harbor constructs such as described above. Another aspect of the invention relates to methods for modulating expression of specific genes in transgenic plants by expression of the structural gene component of the constructs, by regulation of expression of one or more endogenous genes in a transgenic plant or by suppression of expression of the polynucleotides of the invention in a transgenic plant. Methods of modulation of gene expression include without limitation (1) inserting into a host cell additional copies of a polynucleotide comprising a coding sequence; (2) modulating an endogenous promoter in a host cell; (3) inserting antisense

4

or ribozyme constructs into a host cell and (4) inserting into a host cell a polynucleotide comprising a sequence encoding a mutant, fragment, or fusion of the native polypeptides of the instant invention.

#### BRIEF DESCRIPTION OF THE TABLES

The sequences of exemplary SDFs and polypeptides encoded thereby of the instant invention are listed in SEQ TABLES 1 and 2; annotation relevant to the sequences shown in SEQ TABLES 1 and 2 is presented in REF TABLES 1 and 2. Each sequence corresponds to a Maximum Length cDNA Polynucleotide Sequence. Each Maximum Length cDNA Polynucleotide Sequence (Ac) corresponds to at least one sequence in SEQ TABLE 1 and 2. REF TABLE 1 corresponds with SEQ TABLE 1; REF TABLE 2 corresponds with SEQ TABLE 2.

REF TABLES 1 and 2 are Reference Tables which correlate each of the sequences and SEQ ID NOS in SEQ TABLES 1 and 2 with a corresponding Maximum Length cDNA Sequence (Ac), Ceres (Applicant) sequence identifier and other information about the individual sequence. SEQ TABLES 1 and 2 are Sequence Tables containing the sequence of each nucleic acid and amino acid sequence.

In REF TABLES 1 and 2, each section begins by identifying the Maximum Length cDNA Polynucleotide Sequence, indicating a "Clone ID" that is a number used for identification purposes by the applicant and in some instances a "Public Genomic DNA" sequence, indicated by a "gi No". In those instances where a public sequence is recited, there follows information about gene annotations such as predicted exons. In this portion, after the description of the gene's functional unit, the starting and ending nucleotide number of the public sequence and the computer program used to generate the result are listed. "INTR" denotes an initial exon. "INTR" denotes an internal exon.

5

"TERM" denotes a terminal exon. For those sequences where the computer program is identified as "OCKHAM", "INIT" denotes the first detected exon and "TERM" denotes the last detected exon. "SINGLE" denotes the entire mRNA coding portion is in one exon. Other notations are considered intuitive to the skilled artisan.

In the next section (Ac), the cDNA MLS is identified by its SEQ ID NO ("Pat. Appln. SEQ ID NO") and the Ceres sequence identifier ("Ceres seq\_id"), which is also merely an identifier useful for the applicant. The designation of "Alternative transcription start sites" can include both positive and negative numbers. Positive numbers refer to the referenced SEQ ID NO directly. The positions indicated by negative numbers, if any, refer to positions in the public genomic sequences. In instances where there is a "Public Genomic DNA" referenced, the relevant genomic sequence can be found by direct reference to the nucleotide sequence indicated by the "gi" number shown for the Public Genomic DNA. In instances where there is no Public Genomic DNA referenced, the relevant nucleotide sequence for alignment is the nucleotide sequence associated with the amino acid sequence designated by a "gi" number in the section (Dp). In these instances, the nucleotide sequence is found in GENBANK by clicking on the link in the National Center for Biotechnology Information Entrez database. The numbering is relative to position 1 as determined by aligning the first residue of the MLS cDNA sequence (SEQ ID NO \*) with the genomic sequence corresponding to the relevant "gi" number.

Subsection (B) lists SEQ ID NOS and Ceres seqs for polypeptide sequences encoded by the cDNA sequence and the location of the start codon within the cDNA sequence that codes for the polypeptide. Subsection (B) also describes additional features within the polypeptide sequence.

6

Subsection (C) provides information regarding identified domains (where present) within the polypeptide and (where present) a name for the polypeptide. Subsection (Dp) provides (where present) information concerning amino acid sequences that are found to be related and have some sequence identity to the polypeptide sequences of SEQ TABLES 1 and 2. Those "related" sequences are identified by a "gi" number and are amino acid sequences in the publicly accessible BLAST databases on the NCBI FTP web site (accessible at [ncbi.nlm.gov/blast](http://ncbi.nlm.gov/blast)). The database at the NCBI FTP site utilizes the "gi" identifiers to assign by NCBI a unique identifier for each sequence in the databases, thereby providing a non-redundant database for sequences from various databases, including GenBank, EMBL, DDBJ (DNA Database of Japan) and PDB (Brookhaven Protein Data Bank). Subsection (Ba), when present, describes a sequence as being considered plant-specific (i.e. a gene found only in a plant) or describes a biochemical activity for the protein encoded by the exemplary SDF. Subsection (Dn) provides polynucleotide sequences (where present) related to the Maximum Length cDNA sequence.

#### DETAILED DESCRIPTION OF THE INVENTION

The invention relates to (I) polynucleotides and methods of use thereof, such as

- 25 IA. Probes, Primers and Substrates;
- IB. Methods of Detection and Isolation;
  - B.1. Hybridization;
  - B.2. Methods of Mapping;
  - B.3. Southern Blotting;
  - B.4. Isolating cDNA from Related Organisms;
  - B.5. Isolating and/or Identifying Orthologous Genes
- 30 IC. Methods of Inhibiting Gene Expression
  - C.1. Antisense

- C.2. Ribozyme Constructs;
- C.3. Co-Suppression;
- C.4. Other Methods to Inhibit Gene Expression
- ID. Methods of Functional Analysis;
- IE. Promoter Sequences and Their Use;
- IF. UTR's and/or Intron Sequences and Their Use; and
- IG. Coding Sequences and Their Use.

The specification also discloses (II) polypeptides including, without limitation, native proteins, mutants, fragments, and fusions. Antibodies to said polypeptides are also disclosed.

The specification also discloses (III) methods of modulating polypeptide production or activity. Examples of such methods include (i) suppressed, (ii) enhanced, and (iii) directed expression.

The specification also discloses (IV) gene constructs and construction of expression vectors, including (IVA) coding sequences and (IVB) promoters, and (IVC) Signal Peptides, (V) transformation: procedures to illustrate the invention by way of examples.

#### I. Polynucleotides

A number of the nucleotide sequences disclosed in SEQ TABLES 1 AND 2 herein as representative of the SDFs of the invention can be obtained by sequencing genomic DNA (gDNA) and/or cDNA from corn plants grown from HYBRID SEED # 35A19, purchased from Pioneer Hi-Bred International, Inc., Supply Management, P.O. Box 256, Johnston, Iowa 50131-0256.

Exemplified SDFs of the invention represent portions of the genome of corn or *Arabidopsis* and/or represent mRNA expressed from that genome. The isolated nucleic acid of the invention also encompasses corresponding portions of the genome

and/or cDNA complement of other organisms as described in detail below.

Starting material for cDNA synthesis for the exemplary corn cDNA clones having sequences presented in SEQ TABLES 1 AND 2 was poly(A)-containing polysomal mRNAs from inflorescences and root tissues of corn plants grown from HYBRID SEED # 35A19. Male inflorescences and female (pre- and post-fertilization) inflorescences were isolated at various stages of development. Selection for poly(A) containing polysomal RNA was done using oligo d(T) cellulose columns, as described by Cox and Goldberg, "Plant Molecular Biology: A Practical Approach", pp. 1-35, Shaw ed., c. 1988 by IRL, Oxford.

Tissues were or each organ was individually pulverized and frozen in liquid nitrogen. Next, the samples were homogenized in the presence of detergents and then centrifuged. The debris and nuclei were removed from the sample and more detergents were added to the sample. The sample was centrifuged and the debris was removed. Then the sample was applied to a 2M sucrose cushion to isolate polysomes. The RNA was isolated by treatment with detergents and proteinase K followed by ethanol precipitation and centrifugation. The polysomal RNA from the different tissues was pooled according to the following mass ratios: 15/15/1 for male inflorescences, female inflorescences and root, respectively. The pooled material was then used for cDNA synthesis by the methods described below.

The quality and the integrity of the poly(A) RNAs were evaluated.

A number of the nucleotide sequences disclosed in SEQ TABLES 1 AND 2 herein as representative of the SDFs of the invention can also be obtained by sequencing genomic DNA from *Arabidopsis thaliana*, Wassilewskija ecotype or by sequencing cDNA obtained from mRNA from such plants as described below.

This is a true breeding strain. Seeds of the plant are available from the Arabidopsis Biological Resource Center at the Ohio State University, under the accession number CS2360. Seeds of this plant were deposited under the terms and conditions of the Budapest Treaty at the American Type Culture Collection, Manassas, VA on August 31, 1999, and were assigned ATCC No. ZTA-595.

Starting material for cDNA synthesis for the exemplary *Arabidopsis* cDNA clones having sequences presented in SEQ TABLES 1 AND 2 was polysomal RNA isolated from the top-most inflorescence tissues and roots of *Arabidopsis thaliana* Landsberg erecta (L. er.) also obtained from the Arabidopsis Biological Resource Center. Nine parts inflorescence to every part root was used, as measured by mass. Tissue was pulverized and exposed to liquid nitrogen. Next, the sample was homogenized in the presence of detergents and then centrifuged. The debris and nuclei were removed from the sample and more detergents were added to the sample. The sample was centrifuged and the debris was removed and the sample was applied to a 2M sucrose cushion to isolate polysomal RNA. Cox et al., "Plant Molecular Biology: A Practical Approach", pp. 1-35, Shaw ed., c. 1988 by IRL, Oxford. The polysomal RNA was used for cDNA synthesis by the methods described below. Polysomal mRNA was then isolated as described above for corn cDNA. The quality of the RNA was assessed electrophoretically.

Following preparation of the mRNAs from various tissues as described above, selection of mRNA with intact 5' ends and specific attachment of an oligonucleotide tag to the 5' end of such mRNA was performed using either a chemical or enzymatic approach. Both techniques take advantage of the presence of the "cap" structure, which characterizes the 5' end of most intact mRNAs and which comprises a guanosine generally methylated once, at the 7 position.

The chemical modification approach involves the optional elimination of the 2', 3'-cis diol of the 3' terminal ribose, the oxidation of the 2', 3'-cis diol of the ribose linked to the cap of the 5' ends of the mRNAs into a dialdehyde, and the coupling of the such obtained dialdehyde to a derivatized oligonucleotide tag. Further detail regarding the chemical approaches for obtaining mRNAs having intact 5' ends are disclosed in International Application No. WO96/34981 published November 7, 1996.

The enzymatic approach for ligating the oligonucleotide tag to the intact 5' ends of mRNAs involves the removal of the phosphate groups present on the 5' ends of uncapped incomplete mRNAs, the subsequent decapping of mRNAs having intact 5' ends and the ligation of the phosphate present at the 5' end of the decapped mRNA to an oligonucleotide tag. Further detail regarding the enzymatic approaches for obtaining mRNAs having intact 5' ends are disclosed in Dumas Milne Edwards J.B. (Doctoral Thesis of Paris VI University, Le clonage des ADNC complets: difficultes et perspectives nouvelles. Apports pour l'etude de la regulation de l'expression de la tryptophane hydroxylase de rat, 20 Dec. 1993), EP0 625572 and Kato et al., Gene 150:243-250 (1994).

In both the chemical and the enzymatic approach, the oligonucleotide tag has a restriction enzyme site (e.g. an EcoRI site) therein to facilitate later cloning procedures. Following attachment of the oligonucleotide tag to the mRNA, the integrity of the mRNA is examined by performing a Northern blot using a probe complementary to the oligonucleotide tag.

For the mRNAs joined to oligonucleotide tags using either the chemical or the enzymatic method, first strand cDNA synthesis is performed using an oligo-dT primer with reverse transcriptase. This oligo-dT primer can contain an internal tag of at least 4 nucleotides, which can be different from one mRNA preparation to another. Methylated dCTP is used for cDNA

11

first strand synthesis to protect the internal EcoRI sites from digestion during subsequent steps. The first strand cDNA is precipitated using isopropanol after removal of RNA by alkaline hydrolysis to eliminate residual primers.

5 Second strand cDNA synthesis is conducted using a DNA polymerase, such as Klenow fragment and a primer corresponding to the 5' end of the ligated oligonucleotide. The primer is typically 20-25 bases in length. Methylated dCTP is used for second strand synthesis in order to protect internal EcoRI sites in the cDNA from digestion during the cloning process.

10 Following second strand synthesis, the full-length cDNAs are cloned into a phagemid vector, such as pBlueScript™ (Stratagene). The ends of the full-length cDNAs are blunted with T4 DNA polymerase (Biolabs) and the cDNA is digested with EcoRI. Since methylated dCTP is used during cDNA synthesis, the EcoRI site present in the tag is the only hemi-methylated site; hence the only site susceptible to EcoRI digestion. In some instances, to facilitate subcloning, an Hind III adapter is added to the 3' end of full-length cDNAs.

20 The full-length cDNAs are then size fractionated using either exclusion chromatography (Acs, Biosepra) or electrophoretic separation which yields 3 to 6 different fractions. The full-length cDNAs are then directionally cloned either into pBlueScript™ using either the EcoRI and SmaI restriction sites or, when the Hind III adapter is present in the full-length cDNAs, the EcoRI and Hind III restriction sites. The ligation mixture is transformed, preferably by electroporation, into bacteria, which are then propagated under appropriate antibiotic selection.

30 Clones containing the oligonucleotide tag attached to full-length cDNAs are selected as follows.

The plasmid cDNA libraries made as described above are purified (e.g. by a column available from Qiagen). A positive selection of the tagged clones is performed as follows.

12

Briefly, in this selection procedure, the plasmid DNA is converted to single stranded DNA using phage F1 gene II endonuclease in combination with an exonuclease (Chang et al., Gene 127:95 (1993)) such as exonuclease III or T7 gene 6 exonuclease. The resulting single stranded DNA is then

5 purified using paramagnetic beads as described by Fry et al., Biotechniques 13: 124 (1992). Here the single stranded DNA is hybridized with a biotinylated oligonucleotide having a sequence corresponding to the 3' end of the oligonucleotide tag. Preferably, the primer has a length of 20-25 bases. Clones including a sequence complementary to the biotinylated oligonucleotide are selected by incubation with streptavidin coated magnetic beads followed by magnetic capture. After capture of the positive clones, the plasmid DNA is released from the magnetic beads and converted into double stranded DNA using a DNA polymerase such as ThermoSequenase™ (obtained from

15 Amersham Pharmacia Biotech). Alternatively, protocols such as the Gene Trapper™ kit (Gibco BRL) can be used. The double stranded DNA is then transformed, preferably by electroporation, into bacteria. The percentage of positive clones having the 5' tag oligonucleotide is typically estimated to be between 90 and 98% from dot blot analysis.

20 Following transformation, the libraries are ordered in microtiter plates and sequenced. The Arabidopsis library was deposited at the American Type Culture Collection on January 7, 2000 as "E-coli liba 010600" under the accession number \_\_\_\_\_.

Other methods for cloning full-length cDNA are described, for example, by Seki et al., Plant Journal 15:707-720 (1998) "High-efficiency cloning of Arabidopsis full-length cDNA by biotinylated Cap trapper"; Maruyama et al., Gene 138:171 (1994) "Oligo-capping a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides"; and WO 96/34981.

25

13

It is contemplated that the nucleotide sequences presented herein may contain some small percentage of errors. These errors may arise in the normal course of determination of nucleotide sequences. Sequence errors can be corrected by obtaining seeds deposited under the accession numbers cited above, propagating them, isolating genomic DNA or appropriate mRNA from the resulting plants or seeds thereof, amplifying the relevant portion of the genomic DNA or mRNA using primers having a sequence that flanks the erroneous sequence, and sequencing the amplification product.

10

#### 1.A. Probes, Primers and Substrates

SDFs of the invention can be applied to substrates for use in array applications such as, but not limited to, assays of global gene expression, for example under varying conditions of development, growth conditions. The arrays can also be used in diagnostic or forensic methods.

Probes and primers of the instant invention will hybridize to a polynucleotide comprising a sequence in SEQ TABLES 1 AND 2. Though many different nucleotide sequences can encode an amino acid sequence, in some instances, the sequences of SEQ TABLES 1 AND 2 are preferred for encoding polypeptides of the invention. However, the sequence of the probes and/or primers of the instant invention need not be identical to those in SEQ TABLES 1 AND 2 or the complements thereof. For example, some variation in probe or primer sequence and/or length can allow additional family members to be detected, as well as orthologous genes and more taxonomically distant related sequences. Similarly probes and/or primers of the invention can include additional nucleotides that serve as a label for detecting the formed duplex or for subsequent cloning purposes.

Probe length will vary depending on the application. For use as PCR primers, probes should be 12-40 nucleotides,

14

preferably 18-30 nucleotides long. For use in mapping, probes should be 50 to 500 nucleotides, preferably 100-250 nucleotides long. For Southern hybridizations, probes as long as several kilobases can be used as explained below.

The probes and/or primers can be produced by synthetic procedures such as the triester method of Matteucci et al. *J. Am. Chem. Soc.* 103:3185 (1981); or according to Urdea et al. *Proc. Natl. Acad.* 80:7461 (1981) or using commercially available automated oligonucleotide synthesizers.

10

#### I.B. Methods of Detection and Isolation

##### B.1. Hybridization

Probes and/or primers can be used for detection and/or isolation of polynucleotide sequences. Such polynucleotides are included in the subject matter of the invention. Depending on the stringency of the conditions under which these probes and/or primers are used, polynucleotides exhibiting a wide range of similarity to those in SEQ TABLES 1 AND 2 can be detected or isolated.

"Stringency" is a function of probe length, probe composition (G + C content), and salt concentration, organic solvent concentration, and temperature of hybridization or wash conditions. Stringency is typically compared by the parameter " $T_m$ ", which is the temperature at which 50% of the complementary molecules in the hybridization are hybridized, in terms of a temperature differential from  $T_m$ . High stringency conditions are those providing a condition of  $T_m - 5^\circ\text{C}$  to  $T_m - 10^\circ\text{C}$ . Medium stringency conditions are those providing  $T_m - 20^\circ\text{C}$  to  $T_m - 29^\circ\text{C}$ . Low stringency conditions are those providing a condition of  $T_m - 40^\circ\text{C}$  to  $T_m - 49^\circ\text{C}$ . The relationship of hybridization conditions to  $T_m$  (in  $^\circ\text{C}$ ) is expressed in the mathematical equation

15

$$T_s = 81.5 - 16.6(\log_{10}[\text{Na}^+]) + 0.41(\%G+C) - (600/N) \quad (1)$$

where N is the length of the probe. This equation works well for probes 14 to 70 nucleotides in length that are identical to the target sequence. The equation below for  $T_m$  of DNA-DNA hybrids is useful for probes in the range of 50 to greater than 500 nucleotides, and for conditions that include an organic solvent (formamide).

$$T_m = 81.5 + 16.6 \log_{10} \left( \frac{[\text{Na}^+]}{(1+2.7[\text{M}^+])} \right) + 0.41(\%G+C) - 500/L + 0.63(\% \text{formamide}) \quad (2)$$

where L is the length of the probe in the hybrid. (P. Tijessen, "Hybridization with Nucleic Acid Probes" in Laboratory Techniques in Biochemistry and Molecular Biology, P.C. van der Vliet, ed., c. 1993 by Elsevier, Amsterdam.) The  $T_m$  of equation (2) is affected by the nature of the hybrid; for DNA-RNA hybrids  $T_m$  is 10-15°C higher than calculated, for RNA-RNA hybrids  $T_m$  is 20-25°C higher. Because the  $T_m$  decreases about 1 °C for each 1% decrease in homology when a long probe is used (Bonner et al., *J. Mol. Biol.* 81:123 (1973)), stringency conditions can be adjusted to favor detection of identical genes or related family members.

Equation (2) is derived assuming equilibrium and therefore, hybridizations according to the present invention are most preferably performed under conditions of probe excess and for sufficient time to achieve equilibrium. The time required to reach equilibrium can be shortened by inclusion of a "hybridization accelerator" such as dextran sulfate or another high volume polymer in the hybridization buffer.

When the practitioner wishes to examine the result of membrane hybridizations under a variety of stringencies, an efficient way to do so is to perform the hybridization under a low stringency condition, then to wash the hybridization membrane under increasingly stringent conditions. With

16

respect to wash steps, preferred stringencies lie within the ranges stated above; high stringency is 5-8°C below  $T_m$ , medium stringency is 26-29°C below  $T_m$  and low stringency is 45-48°C below  $T_m$ .

5 A number of methods known to those skilled in the art can be used with the probes and/or primers of the invention to isolate and detect polynucleotides, including, without limitation: Southern, Northern, Branched DNA hybridization assays, polymerase chain reaction, and variations thereof.

10 When using SDFs to identify orthologous genes in other species, the practitioner will preferably adjust the amount of target DNA of each species so that, as nearly as is practical, the same number of genome equivalents are present for each species examined. This prevents faint signals from species having large genomes, and thus small numbers of genome equivalents per mass of DNA, from erroneously being interpreted as absence of the corresponding gene in the genome.

20 A good general discussion of the factors for determining hybridization conditions is provided by Sambrook et al. ("Molecular Cloning, a Laboratory Manual, 2nd ed., c. 1989 by Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; see esp., chapters 11 and 12). Additional considerations and details of the physical chemistry of hybridization are provided by G.H. Keller and M.M. Manak "DNA Probes", 2<sup>nd</sup> Ed. pp. 1-25, c. 1993 by Stockton Press, New York, NY.

Hybridization of one nucleic acid to another constitutes a physical property that defines the subject SDF of the invention. Also, such hybridization imposes structural limitations on the pair. For example, for a probe molecule, given that the sequence of the probe nucleic acid is known and fixed, equation (2) indicates that the combined variation in GC content of the target DNA and mismatch between the

17

probe and the hybridizing DNA is determined for any given hybridization buffer composition and  $T_m$ .

The probes and/or primers of the instant invention can be used to detect or isolate nucleotides that are "identical" to the probes or primers. Two nucleic acid sequences or polypeptides are said to be "identical" if the sequence of nucleotides or amino acid residues, respectively, in the two sequences is the same when aligned for maximum correspondence as described below. The term "complementary to" is used herein to mean that the sequence can form a Watson-Crick base pair with a reference polynucleotide sequence. Complementary sequences can include nucleotides, such as inosine, that neither disrupt Watson-Crick base pairing nor contribute to the pairing.

Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman *Add. APL. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman and Wunsch *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson and Lipman *Proc. Natl. Acad. Sci. (USA)* 85: 2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, PASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection. Given that two sequences have been identified for comparison, GAP and BESTFIT are preferably employed to determine their optimal alignment. Typically, the default values of 5.00 for gap weight and 0.30 for gap weight length are used.

The probes and/or primers of the invention can also be used to detect and/or isolate polynucleotides exhibiting at least 80% sequence identity with the sequences of SEQ TABLES 1 AND 2 or fragments thereof.

"Percentage of sequence identity" is determined by comparing two optimally aligned sequences over a comparison

18

window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (e.g., gaps or overhangs) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity. "Percentage of sequence identity" can be determined by the algorithms described above.

The term "substantial identity" between polynucleotide or polypeptide sequences refers to polynucleotide or polypeptide comprising a sequence that has at least 80% sequence identity, preferably at least 85%, more preferably at least 90% and most preferably at least 95%, even more preferably, at least 96%, 97%, 98% or 99% sequence identity compared to a reference sequence using the programs.

Isolated polynucleotides within the scope of the invention also include allelic variants of the specific sequences presented in SEQ TABLES 1 AND 2. An "allelic variant" is a sequence that is a variant from that of the SDF, but represents the same chromosomal locus in the organism. Allelic variants can arise by normal genetic variation in a population. Allelic variants can also be produced by genetic engineering methods. An allelic variant can be one that is found in a naturally occurring plant, including a cultivar or ecotype. A silent allele can give rise to phenotypic and expression profiles. An allelic variant may or may not give rise to a phenotypic change, and may or may not be expressed. An expressed allele can result in a detectable change in the phenotype of the trait



13

represented by the locus. Allelic variations can occur in any portion of the gene sequence, including regulatory regions as well as structural regions.

With respect to nucleotide sequences, degeneracy of the genetic code provides the possibility to substitute at least one base of the base sequence of a gene with a different base without causing the amino acid sequence of the polypeptide produced from the gene to be changed. Hence, the DNA of the present invention may also have any base sequence that has been changed from a sequence in SEQ TABLES 1 AND 2 by substitution in accordance with degeneracy of genetic code. References describing codon usage include: Careis et al., *J. Mol. Evol.* 46: 45 (1998) and Fennoy et al., *Nucl. Acids Res.* 21(23): 5294 (1993).

15

#### A. B.2. Mapping

The isolated SDF DNA of the invention can be used to create various types of genetic and physical maps of the genome of corn, Arabidopsis or other plants. Some SDFs may be absolutely associated with particular phenotypic traits, allowing construction of gross genetic maps. While not all SDFs will immediately be associated with a phenotype, all SDFs can be used as probes for identifying polymorphisms associated with phenotypes of interest. Briefly, total DNA is isolated from individuals and is subsequently cleaved with one or more restriction enzymes, separated according to mass, transferred to a solid support, hybridized with SDF DNA and the pattern of fragments compared. Polymorphisms associated with a particular SDF are visualized as differences in the size of fragments produced between individual DNA samples after digestion with a particular restriction enzyme and hybridization with the SDF. After identification of polymorphic SDF sequences, linkage studies can be conducted. Recombinants produced are analyzed using the same restriction

20

enzyme/hybridization procedure. After identification of many polymorphisms using SDF sequences, linkage studies can be conducted by using the individuals showing polymorphisms as parents in crossing programs. F2 progeny recombinants or recombinant inbreds, for example, are then analyzed using the same restriction enzyme/hybridization procedure. The order of DNA polymorphisms along the chromosomes can be inferred based on the frequency with which they are inherited together versus independently. The closer two polymorphisms are together in a chromosome the higher the probability that they are inherited together. Integration of the relative positions of all the polymorphisms and associated marker SDFs produces a genetic map of the species, where the distances between markers reflect the recombination frequencies in that chromosome segment.

The use of recombinant inbred lines for such genetic mapping is described for Arabidopsis by Alonso-Blanco et al. (*Methods in Molecular Biology*, vol. 82, "Arabidopsis Protocols", pp. 137-146, J.M. Martinez-Zapater and J. Salinas, eds., c. 1998 by Humana Press, Totowa, NJ) and for corn by Burr ("Mapping Genes with Recombinant Inbreds", pp. 249-254. In Freeling, M. and V. Walbot (Ed.), *The Maize Handbook*, c. 1994 by Springer-Verlag New York, Inc.: New York, NY, USA; Berlin Germany; Burr et al. *Genetics* (1998) 118: 519; Gardiner, J. et al., (1993) *Genetics* 134: 917).

However, this procedure is not limited to plants and can be used for other organisms (such as yeast) or for individual cells.

The SDFs of the present invention can also be used for simple sequence repeat (SSR) mapping. Rice SSR mapping is described by Morgante et al. (*The Plant Journal* (1993) 3: 165), Panaud et al. (*Genome* (1995) 38: 1170); Senior et al. (*Crop Science* (1996) 36: 1676), Taramino et al. (*Genome* (1996) 39: 277) and Ahn et al. (*Molecular and General*

21

Genetics (1993) 241: 483-90). SSR mapping can be achieved using various methods. In one instance, polymorphisms are identified when sequence specific probes flanking an SSR contained within an SDF are made and used in polymerase chain reaction (PCR) assays with template DNA from two or more individuals of interest. Here, a change in the number of tandem repeats between the SSR-flanking sequence produces differently sized fragments (U.S. Patent 5,766,847).

Alternatively, polymorphisms can be identified by using the PCR fragment produced from the SSR-flanking sequence specific primer reaction as a probe against Southern blots representing different individuals (U.S. Refs. et al., (1997) *Electrophoresis* 18: 1519).

Genetic and physical maps of crop species have many uses. For example, these maps can be used to devise positional cloning strategies for isolating novel genes from the mapped crop species. In addition, because the genomes of closely related species are largely syntenic (that is, they display the same ordering of genes within the genome), these maps can be used to isolate novel alleles from wild relatives of crop species by positional cloning strategies.

The various types of maps discussed above can be used with the SDFs of the invention to identify Quantitative Trait Loci (QTLs). Many important crop traits, such as the solids content of tomatoes, are quantitative traits and result from the combined interactions of several genes. These genes reside at different loci in the genome, oftentimes on different chromosomes, and generally exhibit multiple alleles at each locus. The SDFs of the invention can be used to identify QTLs and isolate specific alleles as described by de Vicente and Tanksley (*Genetics* 134:585 (1993)). In addition to isolating QTL alleles present in crop species, the SDFs of the invention can also be used to isolate alleles from the corresponding QTL of wild relatives. Transgenic plants

22

having various combinations of QTL alleles can then be created and the effects of the combinations measured. Once an ideal allele combination has been identified, crop improvement can be accomplished either through biotechnological means or by directed conventional breeding programs (for review see Tanksley and McCouch, *Science* 277:1063 (1997)).

In another embodiment the SDFs can be used to help create physical maps of the genome of corn, *Arabidopsis* and related species. Where SDFs have been ordered on a genetic map, as described above, then SDFs can be used as probes to discover which clones in large libraries of plant DNA fragments in YACs, BACs, etc. contain the same SDF or similar sequences, thereby facilitating the assignment of the large DNA fragments to chromosomal positions. Subsequently, the large BACs, YACs, etc. can be ordered unambiguously by more detailed studies of their sequence composition (e.g. Matra et al. (1997) *Genomic Research* 7:1072-1084) and by using their end or other sequences to find the identical sequences in other cloned DNA fragments. The overlapping of DNA sequences in this way allows large contigs of plant sequences to be built, that, when sufficiently extended, provide a complete physical map of a chromosome. Sometimes the SDFs themselves will provide the means of joining cloned sequences into a contig.

The patent publication W095/35505 and U.S. Patents 5,445,943 and 5,410,270 describe scanning multiple alleles of a plurality of loci using hybridization to arrays of oligonucleotides. These techniques are useful for each of the types of mapping discussed above.

Following the procedures described above and using a plurality of the SDFs of the present invention, any individual can be genotyped. These individual genotypes can be used for the identification of particular cultivars,

23

varieties, lines, ecotypes and genetically modified plants or can serve as tools for subsequent genetic studies involving multiple phenotypic traits.

### B. 3 Southern Blot Hybridization

5

The sequences from SEQ TABLES 1 AND 2 can be used as probes for various hybridization techniques. These techniques are useful for detecting target polynucleotides in a sample or for determining whether transgenic plants, seeds or host cells harbor a gene or sequence of interest and thus might be expected to exhibit a particular trait or phenotype.

In addition, the hybridization of the SDFs of the invention to nucleic acids obtained from other organisms can be used to identify orthologous genes from other species and/or additional members of gene families either in the same or different species. In regard to identifying genes in other species, a Southern blot of genomic DNA provides description of isolated DNA fragments that comprise the orthologous genes or additional members of the gene families. That is, given such data, one of ordinary skill in the art could distinguish the isolated DNA fragments by their size together with the restriction sites at each end and by the property of hybridizing with the SDF probe under the stated conditions.

In addition, the SDFs from the invention can be used to isolate additional members of gene families from the same species and/or orthologous genes from different species. This is accomplished by hybridizing an SDF to a Southern blot containing the appropriate genomic DNA or cDNA. Given the resulting hybridization data, one of ordinary skill in the art could distinguish and isolate the correct DNA fragments by size, restriction sites and stated hybridization conditions from a gel or from a library.

24

Identification and isolation of orthologous genes from closely related species and alleles within a species is particularly desirable because of their potential for crop improvement. Many important crop traits, such as the solid content of tomatoes, result from the combined interactions of the products of several genes residing at different loci in the genome. Generally, alleles at each of these loci can make quantitative differences to the trait. By identifying and isolating numerous alleles for each locus from within or different species, transgenic plants having various combinations of alleles can be created and the effects of the combinations measured. Once a more favorable ideal allele combination has been identified, crop improvement can be accomplished either through biotechnological means or by directed conventional breeding programs (Tanksley et al. Science 277:1063(1997)).

The results from hybridizations of the SDFs of the invention to Southern blots containing DNA from another species can also be used to generate restriction fragment maps for the corresponding genomic regions. These maps provide map provides additional information about the relative positions of restriction sites within fragments, further distinguishing mapped DNA from the remainder of the genome.

Physical maps can be made by digesting genomic DNA with different combinations of restriction enzymes.

Probes for Southern blotting to distinguish individual restriction fragments can range in size from 15 to 20 nucleotides to several thousand nucleotides. More preferably, the probe is 100 to 1000 nucleotides long for identifying members of a gene family when it is found that repetitive sequences would complicate the hybridization. For identifying an entire corresponding gene in another species, the probe is more preferably the length of the gene,

25

typically 2000 to 10,000 nucleotides, but probes 50-1,000 nucleotides long might be used. Some genes, however, might require probes up to 15,000 nucleotides long or overlapping probes constituting the full-length sequence to span their lengths.

Also, while it is preferred that the probe be homogeneous with respect to its sequence, that is not necessary. For example, as described below, a probe representing members of a gene family having diverse sequences can be generated using PCR to amplify genomic DNA or RNA templates using primers derived from SDFs that include sequences that define the gene family.

For identifying corresponding genes in another species, the probe for Southern blotting most preferably would be the genomic copy of the probe gene. This allows all elements of the gene to be identified in the other species. The next most preferable probe is a cDNA spanning the entire coding sequence, which allows all of the mRNA-coding portion of the gene to be identified; in this case it is possible that some introns in the gene might be missed. Probes for Southern blotting can easily be generated from SDFs by making primers having the sequence at the ends of the SDF and using corn or *Arabidopsis* genomic DNA as a template. In instances where the SDF includes sequence conserved among species, primers including the conserved sequence can be used for PCR with genomic DNA from a species of interest to obtain a probe. Similarly, if the SDF includes a domain of interest, that portion of the SDF can be used to make primers and, with appropriate template DNA, used to make a probe to identify genes containing the domain. Alternatively, the PCR products can be resolved, for example by gel electrophoresis, and cloned and/or sequenced. In this manner, the variants of the domain among members of a gene family, both within and across species, can be examined.

26

#### B.4.1 Isolating DNA from Related Organisms

The SDFs of the invention can be used to isolate the corresponding DNA from other organisms. Either cDNA or genomic DNA can be isolated. For isolating genomic DNA, a lambda, cosmid, BAC or YAC, or other large insert genomic library from the plant of interest can be constructed using standard molecular biology techniques as described in detail by Sambrook et al. 1989 (Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed. Cold Spring Harbor Laboratory Press, New York) and by Ausubel et al. 1992 (Current Protocols in Molecular Biology, Greene Publishing, New York).

To screen a phage library, recombinant lambda clones are plated out on appropriate bacterial medium using an appropriate *E. coli* host strain. The resulting plaques are lifted from the plates using nylon or nitrocellulose filters. The plaque lifts are processed through denaturation, neutralization, and washing treatments following the standard protocols outlined by Ausubel et al. (1992). The plaque lifts are hybridized to either radioactively labeled or non-radioactively labeled SDF DNA at room temperature for about 16 hours, usually in the presence of 50% formamide and 5X SSC (sodium chloride and sodium citrate) buffer and blocking reagents. The plaque lifts are then washed at 42°C with 1% Sodium Dodecyl Sulfate (SDS) and at a particular concentration of SSC. The SSC concentration used is dependent upon the stringency at which hybridization occurred in the initial Southern blot analysis performed. For example, if a fragment hybridized under medium stringency (e.g.,  $T_m - 20^\circ\text{C}$ ), then this condition is maintained or preferably adjusted to a less stringent condition (e.g.,  $T_m - 30^\circ\text{C}$ ) to wash the plaque lifts. Positive clones show detectable hybridization e.g., by exposure to X-ray films or chromogen formation. The positive clones are then

27

subsequently isolated for purification using the same general protocol outlined above. Once the clone is purified, restriction analysis can be conducted to narrow the region corresponding to the gene of interest. The restriction analysis and succeeding subcloning steps can be done using procedures described by, for example Sambrook et al. (1989) cited above.

To screen a YAC library, the procedures outlined for the lambda library are essentially similar except the YAC clones are harbored in bacterial colonies. The YAC clones are plated out at reasonable density on nitrocellulose or nylon filters supported by appropriate bacterial medium in petri plates. Following the growth of the bacterial clones, the filters are processed through the denaturation, neutralization, and washing steps following the procedures of Ausubel et al. 1992. The same hybridization procedures for lambda library screening are followed.

To isolate cDNA, similar procedures using appropriately modified vectors are employed. For instance, the library can be constructed in a lambda vector appropriate for cloning cDNA such as  $\lambda$ gt11. Alternatively, the cDNA library can be made in a plasmid vector. cDNA for cloning can be prepared by any of the methods known in the art, but is preferably prepared as described above. Preferably, a cDNA library will include a high proportion of full-length clones.

#### B. 5. Isolating and/or Identifying Orthologous Genes

Probes and primers of the invention can be used to identify and/or isolate polynucleotides related to those in SEQ TABLES 1 AND 2. Related polynucleotides are those that are native to other plant organisms and exhibit either similar sequence or encode polypeptides with similar biological activity. One specific example is an orthologous gene, a gene that has a high degree of sequence similarity, often along the

28

entire length of the coding portion of the gene, and also encodes a gene product that performs a similar function in the organism. Orthologous genes may be distinguished from homologous genes in that homologous genes share sequence similarity but often only in a portion of the sequence, which often represents a functional domain such as a tyrosine kinase activity, a DNA binding domain, or the like. The functional activities of homologous genes are not necessarily the same, but are the same for orthologous genes. The degree of identity is a function of evolutionary separation and, in closely related species, the degree of identity can be 98 to 100%. The amino acid sequence of a protein encoded by an orthologous gene can be less than 75% identical, but tends to be at least 75% or at least 80% identical, more preferably at least 90%, most preferably at least 95% identical to the amino acid sequence of the reference protein.

To find orthologous genes, the probes are hybridized to nucleic acids from a species of interest under low stringency conditions and blots are then washed under conditions of increasing stringency. It is preferable that the wash stringency be such that sequences that are 85 to 100% identical will hybridize. More preferably, sequences 90 to 100% identical will hybridize and most preferably only sequences greater than 95% identical will hybridize. The low stringency condition is preferably one where sequences containing as much as 40-45% mismatches will be able to hybridize. This condition is established by  $T_m - 40^\circ\text{C}$  to  $T_m - 48^\circ\text{C}$  (see below). One of ordinary skill in the art will recognize that, due to degeneracy in the genetic code, amino acid sequences that are identical can be encoded by DNA sequences as little as 67% identical. Thus, it is preferable to make an overlapping series of shorter probes, on the order of 24 to 45 nucleotides, and individually hybridize them to the same arrayed library to

29

avoid the problem of degeneracy introducing large numbers of mismatches.

As evolutionary divergence increases, genome sequences also tend to diverge. Thus, one of skill will recognize that searches for orthologous genes between more divergent species will require the use of lower stringency conditions compared to searches between closely related species. Also, degeneracy is more of a problem for searches in the genome of a species more distant evolutionarily from the species that is the source of the SDF probe sequences.

Therefore the method described in Bouckaert et al., U.S. Ser. No. 60/121,700 Atty. Dkt. No. 2750-117P, Client Dkt. No. 00010.001, filed February 25, 1999, hereby incorporated in its entirety by reference, can be applied to the SDFs of the present invention to isolate related genes from plant species which do not hybridize to the corn or *Arabidopsis* sequences of SEQ TABLES 1 AND 2.

Identification of the relationship of nucleotide or

amino acid sequences among plant species can be done by comparison of the subject nucleotide or amino acid sequence to the sequences of SDFs of the present application presented in SEQ TABLES 1 and 2.

The SDFs of the invention can also be used as probes to search for genes that are related to the SDF within a species. Such related genes are typically considered to be members of a "gene family." In such a case, the sequence similarity will often be concentrated into one or a few portions of the sequence. The portions of similar sequence that define the gene family typically encode a portion of a protein or RNA that has an enzymatic or structural function. The degree of identity in the amino acid sequence of the domain that defines the gene family is preferably at least 70%, more preferably 80 to 95%, most preferably 85 to 99%. To search for members of a gene family within a species, a

30

"low stringency" hybridization is usually performed, but this will depend upon the size, distribution and degree of sequence divergence of domains that define the gene family. SDFs encompassing regulatory regions can be used to identify "coordinately expressed" genes by using the regulatory region portion of the SDF as a probe.

In the instances where the SDFs are identified as being expressed from genes that confer a particular phenotype, then the SDFs can also be used as probes to assay plants of different species for those phenotypes.

#### I.C. Methods to Inhibit Gene Expression

In some instances it is desirable to suppress expression of an endogenous or exogenous gene. A well-known instance is the FLAVOR-SAVOR™ tomato, in which the gene encoding ACC synthase is inactivated by an antisense approach, thus delaying softening of the fruit after ripening. See for example, U.S. Patent No. 5,859,330; U.S. Patent No. 5,723,766; Oeller, et al, Science, 254:437-439(1991); and Hamilton et al, Nature, 346:284-287 (1990). Also, timing of flowering can be controlled by suppression of the *FLOWERING LOCUS C*; high levels of this transcript are associated with late flowering, while absence of *FLC* is associated with early flowering (S.D. Michaels et al., Plant Cell 11:949 (1999)). Also, the transition of apical meristem from production of leaves with associated shoots to flowering is regulated by *TERMINAL FLOWER1*, *APETALAI* and *LEAFY*. Thus, when it is desired to induce a transition from shoot production to flowering, it is desirable to suppress *TFL* expression (S.J. Liljegren, Plant Cell 11:1007 (1999)). As another instance, it has been found that suppression of the ethylene forming enzyme results in arrested ovule development and female sterility that can be reversed by application of ethylene (D. De Martinis et al., Plant Cell 11:1061 (1999)). The ability

31

to manipulate fertility of female plants is useful in the increasing fruit production and creating hybrids.

In an instance when it is desired to express a dominant negative mutation, it often helpful to suppress expression of the endogenous, native protein. Expression of dominant negative mutant proteins is a useful tool for research, for example when a dominant negative mutation of a receptor is used to constitutively activate or suppress a signal transduction cascade, allowing examination of the phenotype and thus the trait(s) controlled by that receptor and pathway.

#### C.1 Antisense

In the case of polynucleotides used to inhibit expression of an endogenous gene, the introduced sequence need not be perfectly identical to a sequence of the target endogenous gene. The introduced polynucleotide sequence will typically be at least substantially identical (as determined above) to the target endogenous sequence.

Some polynucleotide SDFs in SEQ TABLES 1 AND 2 represent sequences that are expressed in corn and/or *Arabidopsis*. Thus the invention includes using these sequences to generate antisense constructs to inhibit transcription and/or translation of said SDFs, typically in a plant cell.

To accomplish this, a polynucleotide segment from the desired gene that can hybridize to the mRNA expressed from the desired gene (the "antisense segment") is operably linked to a promoter such that the antisense strand of RNA will be transcribed when the construct is present in a host cell. A regulated promoter can be used in the construct to control transcription of the antisense segment so that transcription occurs only under desired circumstances.

The antisense segment to be introduced generally will be substantially identical to at least a portion of the endogenous gene or genes to be repressed. The sequence, however, need not

32

be perfectly identical to inhibit expression. Further, the antisense product may hybridize to the untranslated region instead of or in addition to the coding portion of the gene. The vectors of the present invention can be designed such that the inhibitory effect applies to other proteins within a family of genes exhibiting homology or substantial homology to the target gene.

For antisense suppression, the introduced antisense segment sequence also need not be full length relative to either the primary transcription product or fully processed mRNA. Generally, higher sequence identity can be used to compensate for the use of a shorter sequence. Furthermore, the introduced sequence need not have the same intron or exon pattern, and homology of non-coding segments may be equally effective. Normally, a sequence of between about 30 or 40 nucleotides and the full length of the transcript should be used, though a sequence of at least about 100 nucleotides is preferred, a sequence of at least about 200 nucleotides is more preferred, and a sequence of at least about 500 nucleotides is especially preferred.

#### C.2. Ribozymes

It is also contemplated that gene constructs representing ribozymes and based on the SDFs in SEQ TABLES 1 AND 2 are an object of the invention. Ribozymes can also be used to inhibit expression of genes by suppressing the translation of the mRNA into a polypeptide. It is possible to design ribozymes that specifically pair with virtually any target RNA and cleave the phosphodiester backbone at a specific location, thereby functionally inactivating the target RNA. In carrying out this cleavage, the ribozyme is not itself altered, and is thus capable of recycling and cleaving other molecules, making it a true enzyme. The inclusion of ribozyme sequences

33

within antisense RNAs confers RNA-cleaving activity upon them, thereby increasing the activity of the constructs.

A number of classes of ribozymes have been identified. One class of ribozymes is derived from a number of small circular RNAs, which are capable of self-cleavage and replication in plants. The RNAs replicate either alone (viroid RNAs) or with a helper virus (satellite RNAs). Examples include RNAs from avocado sunblotch viroid and the satellite RNAs from tobacco ringspot virus, lucerne transient streak virus, velvet tobacco mottle virus, solanum nodiflorum mottle virus and subterranean clover mottle virus. The design and use of target RNA-specific ribozymes is described in Haselhoff et al. *Nature*, 334:585 (1988).

Like the antisense constructs above, the ribozyme sequence portion necessary for pairing need not be identical to the target nucleotides to be cleaved, nor identical to the sequences in SEQ TABLES 1 AND 2. Generally, the sequence in the ribozyme capable of binding to the target sequence exhibits substantial sequence identity to a sequence in SEQ TABLES 1 AND 2 or the complement thereof, or to a portion of said sequence or complement. Further, the ribozyme sequence also need not be full length relative to either the primary transcription product or fully processed mRNA. The ribozyme can be equally effective in inhibiting mRNA translation by cleaving either in the untranslated or coding regions. Generally, higher sequence identity can be used to compensate for the use of a shorter sequence. Furthermore, the introduced sequence need not have the same intron or exon pattern, and homology of non-coding segments may be equally effective.

### C.3. Sense Suppression

Another method of suppression is by introducing an exogenous copy of the gene to be suppressed. Introduction of

34

expression cassettes in which a nucleic acid is configured in the sense orientation with respect to the promoter into the chromosome of a plant or by a self-replicating virus has been shown to be an effective means by which to induce degradation of mRNAs of target genes. For an example of the use of this method to modulate expression of endogenous genes see, Napoli et al., *The Plant Cell* 2:279 (1990), and U.S. Patents Nos. 5,034,323, 5,231,020, and 5,283,184. Generally, where inhibition of expression is desired, some transcription of the introduced sequence is probably necessary. The effect may occur where the introduced sequence contains no coding sequence per se, but comprises only intron or untranslated sequences homologous to sequences present in the primary transcript of the endogenous sequence. The introduction of only regulatory promoter sequences can also cause interference with the activity of endogenous promoters possessing the same sequence. Thus, the described SDFs can also be used to control transcription. In all of these procedures, the introduced sequence generally will be substantially identical to the endogenous sequence intended to be inactivated. The minimal identity will typically be greater than about 65%, but a higher identity might exert a more effective repression of expression of the endogenous sequences. Sequence identity of more than about 80% is preferred, though about 95% to absolute identity would be most preferred. As with antisense regulation, the effect would likely apply to any other proteins within a similar family of genes exhibiting homology or substantial homology to the suppressing sequence.

### C.4. Other Methods to Inhibit Gene Expression

Yet another means of suppressing gene expression is to insert a polynucleotide into the gene of interest to disrupt transcription or translation of the gene.



35

Low frequency homologous recombination can be used to target a polynucleotide insert to a gene by flanking the polynucleotide insert with sequences that are substantially similar to the gene to be disrupted. Sequences from SEQ TABLES 1 AND 2, fragments thereof, and substantially similar sequence thereto can be used for homologous recombination.

In addition, random insertion of polynucleotides into a host cell genome can also be used to disrupt the gene of interest. Azpiroz-Leehan et al., *Trends in Genetics* 13:152 (1997). In this method, screening for clones from a library containing random insertions is preferred to identifying those that have polynucleotides inserted into the gene of interest. Such screening can be performed using probes and/or primers described above based on sequences from SEQ TABLES 1 AND 2, fragments thereof, and substantially similar sequence thereto. The screening can also be performed by selecting clones or R<sub>1</sub> plants having a desired phenotype.

#### I.D. Methods of Functional Analysis

The constructs described in the methods under I.C. above can be used to determine the function of the polypeptide encoded by the gene that is targeted by the constructs.

Down-regulating the transcription and translation of the targeted gene, the host cell or organisms, such as a plant, may produce phenotypic changes as compared to a wild-type cell or organism. In addition, *in vitro* assays can be used to determine if any biological activity, such as calcium flux, DNA transcription, nucleotide incorporation, etc., are being modulated by the down-regulation of the targeted gene.

Coordinated regulation of sets of genes, e.g., those contributing to a desired polygenic trait, is sometimes necessary to obtain a desired phenotype. SDFs of the invention representing transcription activation and DNA binding domains can be assembled into hybrid transcriptional

36

These hybrid transcriptional activators can be used with their corresponding DNA elements (i.e., those bound by the DNA-binding SDFs) to effect coordinated expression of desired genes (J.J. Schwarz et al., *Mol. Cell. Biol.* 12:266 (1992), A. Martinez et al., *Mol. Gen. Genet.* 261:546 (1999)).

The SDFs of the invention can also be used in the two-hybrid genetic systems to identify networks of protein-protein interactions (L. McAlister-Henn et al., *Methods* 19:330 (1999), J.C. Hu et al., *Methods* 20:80 (2000), M. Golovkin et al., *J. Biol. Chem.* 274:36428 (1999), K. Ichimura et al., *Biochem. Biophys. Res. Comm.* 253:532 (1998)). The SDFs of the invention can also be used in various expression display methods to identify important protein-DNA interactions (e.g. B. Luo et al., *J. Mol. Biol.* 266:479 (1997)).

#### I.E. Promoters

The SDFs of the invention are also useful as structural or regulatory sequences in a construct for modulating the expression of the corresponding gene in a plant or other organism, e.g. a symbiotic bacterium. For example, promoter sequences represented in SEQ TABLES 1 AND 2 can be useful in directing expression of coding sequences either as constitutive promoters or to direct expression in particular cell types, tissues, or organs or in response to environmental stimuli.

The term "promoter" refers to a region of sequence determinants located upstream or downstream from the start of transcription and which are involved in recognition and binding of RNA polymerase and other proteins to initiate transcription. A "plant promoter" is a promoter capable of initiating transcription in plant cells and can be used to drive expression of a translated portion of an SDF. Such promoters need not be of plant origin. For example, promoters derived from plant viruses, such as the CaMV35S promoter or from

37

*Agrobacterium tumefaciens* such as the T-DNA promoters, can be plant promoters. A typical example of a constitutive promoter of plant origin is the promoter of the cowpea trypsin inhibitor gene. Typical examples of temporal and/or tissue specific promoters of plant origin that can be used with the polynucleotides of the present invention, are: PTA29, a promoter which is capable of driving gene expression specifically in tapetum and only during anther development (Koltonow et al., *Plant Cell* 2:1201 (1990); RCC2 and RCC3, promoters that direct root-specific gene expression in rice (Xu et al., *Plant Mol. Biol.* 27:237 (1995); TcbRB27, a root-specific promoter from tobacco (Yamamoto et al., *Plant Cell* 3:371 (1991)).

By "specific promoters" is meant promoters that have a high preference of driving gene expression in the specified tissue and/or at the specified time during the concerned tissue or organ development. By "high preference" is meant at least 3-fold, preferably 5-fold, more preferably at least 10-fold still more preferably at least 20-fold, 50-fold or 100-fold increase in expression in the desired tissue over the expression in any undesired tissue.

A typical example of an inducible promoter, which can be utilized with the polynucleotides of the present invention, is PARSK1, the promoter from the *Arabidopsis* gene encoding a serine-threonine kinase enzyme, and which promoter is induced by dehydration, abscissic acid and sodium chloride (Wang and Goodman, *Plant J.* 8:37 (1995)).

With respect to the SDFs of the present invention a promoter is likely to be a relatively small portion of a genomic DNA (gDNA) sequence located in the first 2000 nucleotides upstream from an initial exon identified in a gDNA sequence or initial "ATG" or methionine codon in a corresponding cDNA or mRNA sequence. Such promoters are more likely to be found in the first 1000 nucleotides upstream of an

38

initial ATG or methionine codon of a cDNA sequence corresponding to a gDNA sequence. In particular, the promoter is usually located upstream of the transcription start site. Such a start site is located at the first exon predicted in the CCKHAM-cDNA predictions. In such an instance, the transcription start site is the first nucleotide of the 5' most exon, if the predictions are in the plus (+) strand, or the 3' most if the predictions are in the minus (-) strand. Alternative transcription start sites may be located between the first nucleotide of the 5' most exon (or the 3' most exon in the minus (-) strand) and the initial ATG or methionine codon in the cDNA sequence. The portions of a particular gDNA sequence that function as a promoter in a plant cell will preferably be found to hybridize at medium or high stringency to gDNA sequences presented in SEQ TABLES 1 AND 2.

Promoters are generally modular in nature. Short DNA sequences representing binding sites for proteins can be separated from each other by intervening sequences of varying length. For example, within a particular functional module protein binding sites may be constituted by regions of 5 to 60, preferably 10 to 30, more preferably 10 to 20 nucleotides. Within such binding sites, there are typically 2 to 6 nucleotides that specifically contact amino acids of the nucleic acid binding protein. The protein binding sites are usually separated from each other by 10 to several hundred nucleotides, typically by 15 to 150 nucleotides, often by 20 to 50 nucleotides. DNA binding sites in promoter elements often display dyad symmetry in their sequence. Often elements binding several different proteins, and/or a plurality of sites that bind the same protein, will be combined in a region of 100 to 1000 basepairs.

Elements that have transcription regulatory function can be isolated from their corresponding endogenous gene, or the desired sequence can be synthesized, and recombined in

39

constructs to direct expression of a structural gene in a desired tissue-specific, temporal-specific or other desired manner of inducibility or suppression. When hybridizations are performed to identify or isolate elements of a promoter by hybridization to the long sequences presented in SEQ TABLES 1 AND 2, conditions should be adjusted to account for the above-described nature of promoters. For example short probes, constituting the element sought, should be used under low temperature and/or high salt conditions. When long probes, which might include several promoter elements are used, low to medium stringency conditions are preferred when hybridizing to promoters across species.

Promoters can consist of a "basal promoter" that functions as a site for assembly of a transcription complex comprising an RNA polymerase, for example RNA polymerase II. A typical transcription complex will include additional factors such as TFIIB, TFIID, and TFIIE. Of these, TFIID appears to be the only one to bind DNA directly. Basal promoters frequently include a "TATA box" element usually located between 20 and 35 nucleotides upstream from the site of initiation of transcription. Basal promoters also sometimes include a "CCAAT box" element (typically a sequence CCAAT) and/or a GGGCG sequence, usually located between 40 and 200 nucleotides, preferably 60 to 120 nucleotides, upstream from the start site of transcription.

The promoter might also contain one or more "enhancers" and/or "suppressors" that function as binding sites for additional transcription factors that have the function of modulating the level of transcription with respect to tissue specificity of transcription, transcriptional responses to particular environmental or nutritional factors, and the like.

If a nucleotide sequence of an SDF, or part of the SDF, functions as a promoter or portion of a promoter, then nucleotide substitutions, insertions or deletions that do not

40

substantially affect the binding of relevant DNA binding proteins would be considered equivalent to the exemplified nucleotide sequence. It is envisioned that there are instances where it is desirable to decrease the binding of relevant DNA binding proteins to "silence" or "down-regulate" a promoter, or conversely to increase the binding of relevant DNA binding proteins to "enhance" or "up-regulate" a promoter. In such instances, polynucleotides representing changes to the nucleotide sequence of the DNA-protein contact region by insertion of additional nucleotides, changes to identity of relevant nucleotides, including use of chemically-modified bases, or deletion of one or more nucleotides are considered encompassed by the present invention.

Promoter function can be assayed by methods known in the art, preferably by measuring activity of a reporter gene operatively linked to the sequence being tested for promoter function. Examples of reporter genes include those encoding luciferase, green fluorescent protein, G.S, neo, cat and bar.

#### 20 I.F. UTRs and Junctions

Polynucleotides comprising untranslated (UTR) sequences and intron/exon junctions are also within the scope of the invention. UTR sequences include introns and 5' or 3' untranslated regions (5' UTRs or 3' UTRs). Portions of the sequences shown in SEQ TABLES 1 AND 2 can comprise UTRs and introns or intron/exon junctions.

These portions of SDFs, especially UTRs, can have regulatory functions related to, for example, translation rate and mRNA stability. Thus, these portions of SDFs can be isolated for use as elements of gene constructs for expression of polynucleotides encoding desired polypeptides.

Introns of genomic DNA segments might also have regulatory functions. Sometimes promoter elements,

41

especially transcription enhancer or suppressor elements, are found within introns. Also, elements related to stability of heteronuclear RNA and efficiency of transport to the cytoplasm for translation can be found in intron elements. Thus, these segments can also find use as elements of expression vectors intended for use to transform plants.

Just as with promoters, introns and UTR sequences and intron/exon junctions can vary from those shown in SEQ TABLES 1 AND 2. Such changes from those sequences preferably will not affect the regulatory activity of the UTRs or intron or intron/exon junction sequences on expression, transcription, or translation. However, in some instances, down-regulation of such activity may be desired to modulate traits or phenotypic or *in vitro* activity.

## 15 I.G. Coding Sequences

Isolated polynucleotides of the invention can include coding sequences that encode polypeptides comprising an amino acid sequence encoded by a sequence in SEQ TABLES 1 AND 2 or an amino acid sequence presented in SEQ TABLES 1 AND 2.

A nucleotide sequence "encodes" a polypeptide if a cell (or a cell free *in vitro* system) expressing that nucleotide sequence produces a polypeptide having the recited amino acid sequence when the nucleotide sequence is transcribed and the primary transcript is subsequently processed and translated by a host cell (or a cell free *in vitro* system) harboring the nucleic acid. Thus, an isolated nucleic acid that "encodes" a particular amino acid sequence can be a genomic sequence comprising exons and introns or a cDNA sequence that represents the product of splicing thereof. An isolated nucleic acid "encoding an amino acid sequence" also encompasses heteronuclear RNA, which contains sequences that are spliced out during expression, and mRNA, which lacks those sequences.

42

Coding sequences can be constructed using chemical synthesis techniques or by isolating coding sequences or by modifying such synthesized or isolated coding sequences as described above.

In addition to encoding the polypeptide sequences of SEQ TABLES 1 AND 2, which are native to corn or *Arabidopsis*, the isolated polynucleotides can be variant polynucleotides that encode mutants, fragments, and fusions of those native proteins. Such polypeptides are described below in part II.

In variant polynucleotides generally, the number of substitutions, deletions or insertions is preferably less than 20%, more preferably less than 15%; even more preferably less than 10%, 5%, 3% or 1% of the number of nucleotides comprising a particularly exemplified sequence. It is generally expected that non-degenerate nucleotide sequence changes that result in 1 to 10, more preferably 1 to 5 and most preferably 1 to 3 amino acid insertions, deletions or substitutions will not greatly affect the function of an encoded polypeptide. The most preferred embodiments are those wherein 1 to 20, preferably 1 to 10, most preferably 1 to 5 nucleotides are added to, deleted from and/or substituted in the sequences specifically disclosed in SEQ TABLES 1 AND 2.

Insertions or deletions in polynucleotides intended to be used for encoding a polypeptide should preserve the reading frame. This consideration is not so important in instances when the polynucleotide is intended to be used as a hybridization probe.

## II. Polypeptides

Polypeptides within the scope of the invention include both native proteins as well as mutants, fragments, and fusions thereof. Polypeptides of the invention are those encoded by any of the six reading frames of sequences shown

in SEQ TABLES 1 AND 2, preferably encoded by the three frames reading in the 5' to 3' direction of the sequences as shown.

Native polypeptides include the proteins encoded by the sequences shown in SEQ TABLES 1 AND 2. Such native polypeptides include those encoded by allelic variants.

5 Variants, including mutants, will exhibit at least 80% sequence identity to those native polypeptides of SEQ TABLES 1 AND 2. Sequence identity is used for polypeptides as defined above for polynucleotides. More preferably, the variants will exhibit at least 85% sequence identity; even more preferably, at least 90% sequence identity; more preferably at least 95%, 96%, 97%, 98%, or 99% sequence identity. "Fragments" of polypeptide or "portions" of polypeptides will exhibit similar degrees of identity to the relevant portions of the native polypeptide. Fusions will exhibit similar degrees of identity in that portion of the fusion represented by the variant of the native peptide.

15 Furthermore, variants will exhibit at least one of the functional properties of the native protein. Such properties include, without limitation, protein interaction, DNA interaction, biological activity, immunological activity, receptor binding, signal transduction, transcription activity, growth factor activity, secondary structure, three-dimensional structure, etc. As to properties related to *in vitro* or *in vivo* activities, the variants preferably exhibit at least 60% of the activity of the native protein; more preferably at least 70%, even more preferably at least 80%, 85%, 90% or 95% of at least one activity of the native protein.

30 A type of mutant of the native polypeptides comprises amino acid substitutions. "Conservative substitutions" are preferred to maintain the function or activity of the polypeptide. Such substitutions include conservation of charge, polarity, hydrophobicity, size, etc. For example, one or more amino acid residues within the sequence can be

substituted with another amino acid of similar polarity that acts as a functional equivalent, for example providing a hydrogen bond in an enzymatic catalysis. Substitutes for an amino acid within an exemplified sequence are preferably made among the members of the class to which the amino acid belongs.

5 For example, the nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan and methionine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, 10 asparagine, and glutamine. The positively charged (basic) amino acids include arginine, lysine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

15 Within the scope of sequence identity described above, a polypeptide of the invention may have additional individual amino acids or amino acid sequences inserted into the polypeptide in the middle thereof and/or at the N-terminal and/or C-terminal ends thereof. Likewise, some of the amino acids or amino acid sequences may be deleted from the 20 polypeptide.

#### Antibodies

Isolated polypeptides can be utilized to produce antibodies. Polypeptides of the invention can generally be used, for example, as antigens for raising antibodies by known techniques. The resulting antibodies are useful as reagents for determining the distribution of the antigen protein within the tissues of a plant or within a cell of a plant. The antibodies are also useful for examining the expression level of proteins in various tissues, for example in a wild-type plant or following genetic manipulation of a plant, by methods such as Western blotting.

Antibodies of the present invention, both polyclonal and monoclonal, may be prepared by conventional methods. In

45

general, the polypeptides of the invention are first used to immunize a suitable animal, such as a mouse, rat, rabbit, or goat. Rabbits and goats are preferred for the preparation of polyclonal sera due to the volume of serum obtainable, and the availability of labeled anti-rabbit and anti-goat antibodies as detection reagents. Immunization is generally performed by mixing or emulsifying the protein in saline, preferably in an adjuvant such as Freund's complete adjuvant, and injecting the mixture or emulsion parenterally (generally subcutaneously or intramuscularly). A dose of 50-200 µg/injection is typically sufficient. Immunization is generally boosted 2-6 weeks later with one or more injections of the protein in saline, preferably using Freund's incomplete adjuvant. One may alternatively generate antibodies by *in vitro* immunization using methods known in the art, which for the purposes of this invention is considered equivalent to *in vivo* immunization.

Polyclonal antisera is obtained by bleeding the immunized animal into a glass or plastic container, incubating the blood at 25°C for one hour, followed by incubating the blood at 4°C for 2-18 hours. The serum is recovered by centrifugation (e.g., 1,000xg for 10 minutes). About 20-50 ml per bleed may be obtained from rabbits.

Monoclonal antibodies are prepared using the method of Kohler and Milstein, *Nature* 256: 495 (1975), or modification thereof. Typically, a mouse or rat is immunized as described above. However, rather than bleeding the animal to extract serum, the spleen (and optionally several large lymph nodes) is removed and dissociated into single cells. If desired, the spleen cells can be screened (after removal of nonspecifically adherent cells) by applying a cell suspension to a plate, or well, coated with the protein antigen. B-cells expressing membrane-bound immunoglobulin specific for the antigen bind to the plate, and are not rinsed away with the rest of the suspension. Resulting B-cells, or all dissociated spleen

46

cells, are then induced to fuse with myeloma cells to form hybridomas, and are cultured in a selective medium (e.g., hypoxanthine, aminopterin, thymidine medium, "HAT"). The resulting hybridomas are plated by limiting dilution, and are assayed for the production of antibodies which bind specifically to the immunizing antigen (and which do not bind to unrelated antigens). The selected Mab-secreting hybridomas are then cultured either *in vitro* (e.g., in tissue culture bottles or hollow fiber reactors), or *in vivo* (as ascites in mice).

Other methods for sustaining antibody-producing B-cell clones, such as by EBV transformation, are known.

If desired, the antibodies (whether polyclonal or monoclonal) may be labeled using conventional techniques.

Suitable labels include fluorophores, chromophores, radioactive atoms (particularly <sup>32</sup>P and <sup>125</sup>I), electron-dense reagents, enzymes, and ligands having specific binding partners. Enzymes are typically detected by their activity. For example, horseradish peroxidase is usually detected by its ability to convert 3,3',5,5'-tetramethylbenzidine (TNB) to a blue pigment, quantifiable with a spectrophotometer. "Specific binding partner" refers to a protein capable of binding a ligand molecule with high specificity, as for example in the case of an antigen and a monoclonal antibody specific therefor. Other

specific binding partners include biotin and avidin or streptavidin, IgG and protein A, and the numerous receptor-ligand couples known in the art. It should be understood that the above description is not meant to categorize the various labels into distinct modes. For example, <sup>125</sup>I may serve as a radioactive label or as an electron-dense reagent. HRP may serve as an enzyme or as an antigen for a Mab. Further one may combine various labels for desired effect. For example, Mabs and avidin also require labels in the practice of this invention: thus, one might label a Mab with biotin, and detect

47

its presence with avidin labeled with <sup>125</sup>I, or with an anti-biotin Mab labeled with HRP. Other permutations and possibilities will be readily apparent to those of ordinary skill in the art, and are considered as equivalents within the scope of the instant invention.

#### In Vitro Applications of Polypeptides

Some polypeptides of the invention will have enzymatic activities that are useful *in vitro*. For example, the soybean trypsin inhibitor (Kunitz) family is one of the numerous families of proteinase inhibitors. It comprises plant proteins which have inhibitory activity against serine proteinases from the trypsin and subtilisin families, thiol proteinases and aspartic proteinases. Thus, these peptides find *in vitro* use in protein purification protocols and perhaps in therapeutic settings requiring topical application of protease inhibitors.

Delta-aminolevulinic acid dehydratase (EC 4.2.1.24) (ALAD) catalyzes the second step in the biosynthesis of heme, the condensation of two molecules of 5-aminolevulinic acid to form porphobilinogen. Thus, ALAD proteins can be used as catalysts in synthesis of heme derivatives. Enzymes of biosynthetic pathways generally can be used as catalysts for *in vitro* synthesis of the compounds representing products of the pathway.

Polypeptides encoded by SDFs of the invention can be engineered to provide purification reagents to identify and purify additional polypeptides that bind to them. This allows one to identify proteins that function as multimers or elucidate signal transduction or metabolic pathways. In the case of DNA binding proteins, the polypeptide can be used in a similar manner to identify the DNA determinants of specific binding (S. Pierrou et al., *Anal. Biochem.* 229:99 (1995), S.

48

Chusacutanachai et al., *J. Biol. Chem.* 274:23591 (1999), Q. Lin et al., *J. Biol. Chem.* 272:27274 (1997)).

#### II.A. MUTANTS, FRAGMENTS, AND FUSIONS

Generally, mutants, fragments, or fusions of the polypeptides encoded by the maximum length sequence (MLS) can exhibit at least one of the activities of the identified domains and/or related polypeptides described in Sections (C) and (D) of REF TABLES 1 and 2 corresponding to the MLS of interest.

##### II.A.(1) Mutants

A type of mutant of the native polypeptides comprises amino acid substitutions. "Conservative substitutions", described above (see II.), are preferred to maintain the function or activity of the polypeptide.

Such substitutions include conservation of charge, polarity, hydrophobicity, size, etc. For example, one or more amino acid residues within the sequence can be substituted with another amino acid of similar polarity that acts as a functional equivalent, for example providing a hydrogen bond in an enzymatic catalysis. Substitutes for an amino acid within an exemplified sequence are preferably made among the members of the class to which the amino acid belongs.

For example, the nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan and methionine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine. The positively charged (basic) amino acids include arginine, lysine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

Within the scope of sequence identity described above, a polypeptide of the invention may have additional individual

amino acids or amino acid sequences inserted into the polypeptide in the middle thereof and/or at the N-terminal and/or C-terminal ends thereof. Likewise, some of the amino acids or amino acid sequences may be deleted from the polypeptide. Amino acid substitutions may also be made in the sequences; conservative substitutions being preferred.

One preferred class of mutants are those that comprise (1) the domain of a MLS encoded polypeptide and/or (2) residues conserved between the MLS encoded polypeptide and related polypeptides of the MLS. For this class of mutants, the MLS encoded polypeptide sequence is changed by insertion, deletion, or substitution at positions flanking the domain and/or conserved residues.

Another class of mutants includes those that comprise a MLS encoded polypeptide sequence that is changed in the domain or conserved residues by a conservative substitution.

Yet another class of mutants includes those that lack one of the *in vitro* activities, or structural features of the MLS encoded polypeptides. One example is dominant negative mutants. Such a mutant may comprise an MLS encoded polypeptide sequence with non-conservative changes in a particular domain or group of conserved residues.

#### II.A.(2) FRAGMENTS

Fragments of particular interest are those that comprise a domain identified for a polypeptide encoded by an MLS of the instant invention and mutants thereof. Also, fragments that comprise at least one region of residues conserved between an MLS encoded polypeptide and its related polypeptides are of great interest. Fragments are sometimes useful as dominant negative mutations.

#### II.A.(3) FUSIONS

Of interest are chimeras comprising (1) a fragment of the MLS encoded polypeptide or mutants thereof of interest and (2) a fragment of a polypeptide comprising the same domain. For example, an AP2 helix encoded by a MLS of the invention fused to second AP2 helix from ANT protein, which comprises two AP2 helices. The present invention also encompasses fusions of MLS encoded polypeptides, mutants, or fragments thereof fused with related proteins or fragments thereof.

#### 10 DEFINITION OF DOMAINS

The polypeptides of the invention may possess identifying domains as shown in REF TABLES 1 and 2. Domains are fingerprints or signatures that can be used to characterize protein families and/or motifs. Such fingerprints or signatures can comprise conserved (1) primary sequence, (2) secondary structure, and/or (3) three-dimensional conformation. Generally, each domain has been associated with either a family of proteins or a motif. Typically, these families and/or motifs have been correlated with specific *in-vitro* and/or *in-vivo* activities. A domain can be any length, including the entirety of the sequence of a protein. Detailed descriptions of the domains, associated families and motifs, and correlated activities of the polypeptides of the instant invention are described below. Usually, the polypeptides with designated domain(s) can exhibit at least one activity that is exhibited by any polypeptide that comprises the same domain(s).

Specific domains within the MLS encoded polypeptides are indicated by the reference REF TABLES 1 and 2. In addition, the domains within the MLS encoded polypeptide can be defined by the region that exhibits at least 70% sequence identity with the consensus sequences listed in the detailed description below of each of the domains.



51

The majority of the protein domain descriptions given below are obtained from Prosite, (<http://www.expasy.ch/prosite/>), and Pfam, (<http://pfam.wustl.edu/browse.shtml>).

52

# 1. (AAA) AAA-protein family signature

A large family of ATPases has been described [1 to 5] whose key feature is that they share a conserved region of about 220 amino acids that contains an ATP-binding site. This family is now called AAA, for 'A'ATPases 'A' associated with diverse cellular 'A'activities. The proteins that belong to this family either contain one or two AAA domains. Proteins containing two AAA domains:

- Mammalian and drosophila NSF (N-ethylmaleimide-sensitive fusion protein) and the fungal homolog, SEC18. These proteins are involved in intracellular transport between the endoplasmic reticulum and Golgi, as well as between different Golgi cisternae.
- Mammalian transitional endoplasmic reticulum ATPase (previously known as p97 or VCP) which is involved in the transfer of membranes from the endoplasmic reticulum to the golgi apparatus. This protein forms a ring-shaped homooligomer composed of six subunits. The yeast homolog is CDC48 and it may play a role in spindle pole proliferation.
- Yeast protein PAS1, essential for peroxisome assembly and the related protein PAS1 from *Pichia pastoris*.
- Yeast protein AFG2.
- *Sulfolobus acidocaldarius* protein SAV and *Halobacterium salinarum* cdch which may be part of a transduction pathway connecting light to cell division.

Proteins containing a single AAA domain:

- *Escherichia coli* and other bacteria ftsh (or hflB) protein. Ftsh is an ATP-dependent zinc metalloprotease that seems to degrade the heat-shock sigma-32 factor.
- It is an integral membrane protein with a large cytoplasmic C-terminal domain that contain both the AAA and the protease domains.

53

- Yeast protein YME1, a protein important for maintaining the integrity of the mitochondrial compartment. YME1 is also a zinc-dependent protease.

- Yeast protein AFG3 (or YTA10). This protein also seems to contain a AAA domain followed by a zinc-dependent protease domain.

Subunits from the regulatory complex of the 26S proteasome [6] which is involved in the ATP-dependent degradation of ubiquitinated proteins:

10 a) Mammalian subunit 4 and homologs in other higher eukaryotes, in yeast (gene YTA5) and fission yeast (gene mts2).

b) Mammalian subunit 6 (TBP7) and homologs in other higher eukaryotes and in yeast (gene YTA2).

15 c) Mammalian subunit 7 (MSS1) and homologs in other higher eukaryotes and in yeast (gene CIM5 or YTA3).

d) Mammalian subunit 8 (P45) and homologs in other higher eukaryotes and in yeast (SUG1 or CIM3 or TBY1) and fission yeast (gene let1).

20 Other probable subunits such as human TBP1 which seems to influences HIV gene expression by interacting with the virus tat transactivator protein and yeast YTA1 and YTA6.

- Yeast protein BCS1, a mitochondrial protein essential for the expression of the Rieske iron-sulfur protein.

25 - Yeast protein MSP1, a protein involved in intramitochondrial sorting of proteins.

- Yeast protein PAS8, and the corresponding proteins PAS5 from *Pichia pastoris* and PAY4 from *Yarrowia lipolytica*.

30 - Mouse protein SKD1 and its fission yeast homolog (SPAC2G11.06).

- *Caenorhabditis elegans* meiotic spindle formation protein mei-1.

- Yeast protein SAP1.

54

- Yeast protein YTA7.

- *Mycobacterium leprae* hypothetical protein A2126A.

It is proposed that, in general, the AAA domains in these proteins act as ATP- dependent protein clamps [5]. In addition to the ATP-binding 'A' and 'B' motifs, which are located in the N-terminal half of this domain, there is a highly conserved region located in the central part of the domain which was used to develop a signature pattern.

10 Consensus pattern: [LIVMT]-x-[LIVMT]-[LIVMF]-x-[GATMC]-[ST]-[NS]-x(4)-[LIVM]- D-x-A-[LIFA]-x-R

[1] Froehlich K.-U., Fries H.W., Ruediger M., Erdmann R., Botstein D., Mecke D. J. Cell Biol. 114:443-453(1991).

15 [2] Erdmann R., Wiebel F.F., Flessau A., Rytka J., Beyer A., Froehlich K.-U., Kunau W.-H. Cell 64:499-510(1991).

[3] Peters J.-M., Walsh M.J., Franke W.W. EMBO J. 9:1757-1767(1990).

20 [4] Kunau W.-H., Beyer A., Goette K., Marzioch M., Saidowsky J., Skaletz-Rorowski A., Wiebel F.F. Biochimie 75:209-224(1993).

[5] Confalonieri F., Duguet M. BioEssays 17:639-650(1995).

6] Hilt W., Wolf D.H. Trends Biochem. Sci. 21:96-102(1996).

25 2. Aminotransferases class-IV signature

Aminotransferases share certain mechanistic features

with other pyridoxal-phosphate dependent enzymes, such as the covalent binding of the pyridoxal-phosphate group to a lysine residue. On the basis of sequence similarity, these various enzymes can be grouped [1,2] into subfamilies. One of these, called class-IV, currently consists of the following enzymes:

- Branched-chain amino-acid aminotransferase (EC

2.6.1.42) (transaminase B), a bacterial (gene *ilvE*) and eukaryotic enzyme which catalyzes the reversible

55

transfer of an amino group from 4-methyl-2-oxopentanoate to glutamate, to form leucine and 2-oxoglutarate.

- D-alanine aminotransferase (EC 2.6.1.21). A bacterial enzyme which catalyzes the transfer of the amino group from D-alanine (and other D-amino acids) to 2-oxoglutarate, to form pyruvate and D-aspartate.
- 4-amino-4-deoxychorismate (ADC) lyase (gene pabC). A bacterial enzyme that converts ADC into 4-aminobenzoate (PABA) and pyruvate.

The above enzymes are proteins of about 270 to 415

amino-acid residues that share a few regions of sequence similarity. Surprisingly, the best-conserved region does not include the lysine residue to which the pyridoxal-phosphate group is known to be attached, in *ilvE*. The region that has been selected as a signature pattern is located some 40 residues at the C-terminus side of the PIP-lysine

Consensus pattern: E-x-[STAGCI]-x(2)-N-[LIVMFAC]-[FY]-x(6,12)-[LIVNF]-x-T-x(6,8)-[LIVM]-x-[GS]-[LIVM]-x-[KR]-

[1] Green J.M., Merkei W.K., Nichols B.P. J. Bacteriol. 174:5317-5323(1992).

[2] Bairoch A. Unpublished observations (1992).

### 3. Bacterial mutT domain signature

The bacterial mutT protein is involved in the GO system [1] responsible for removing an oxidatively damaged form of guanine (8-hydroxyguanine or 7,8-dihydro-8-oxoguanine) from DNA and the nucleotide pool. 8-oxo-dGTP is inserted opposite to dA and dC residues of template DNA with almost equal efficiency thus leading to A.T to G.C transversions. MutT specifically degrades 8-oxo-dGTP to the monophosphate with the concomitant release of pyrophosphate. MutT is a small

56

protein of about 12 to 15 Kd. It has been shown [2,3] that a region of about 40 amino acid residues, which is found in the N-terminal part of mutT, can also be found in a variety of other prokaryotic, viral, and eukaryotic proteins. These proteins are:

- Streptomyces pneumoniae mutX.
- A mutT homolog from plasmid pSAM2 of Streptomyces ambofaciens.
- Bartonella bacilliformis invasion protein A (gene *invA*).
- Escherichia coli dATP pyrophosphohydrolase.
- Protein D250 from African swine fever viruses.
- Proteins D9 and D10 from a variety of poxviruses.
- Mammalian 7,8-dihydro-8-oxoguanine triphosphatase [EC 3.1.6.-] [4].
- Mammalian diadenosine 5',5'''-P<sub>1</sub>P<sub>4</sub>-tetrakisphosphate asymmetrical hydrolase (Ap4Aase) (EC 3.6.1.17) [5], which cleaves A-5'-PPP-5'A to yield AMP and ATP.
- A protein encoded on the antisense RNA of the basic fibroblast growth factor gene in higher vertebrates.
- Yeast protein YSA1.
- Escherichia coli hypothetical protein yfaO.
- Escherichia coli hypothetical protein ygdU and HI0901, the corresponding Haemophilus influenzae protein.
- Escherichia coli hypothetical protein yjad and HI0432, the corresponding Haemophilus influenzae protein.
- Escherichia coli hypothetical protein yrfE.
- Bacillus subtilis hypothetical protein yqkS.
- Bacillus subtilis hypothetical protein yzqD.
- Yeast hypothetical protein YGL067w.

It is proposed [2] that the conserved domain could be involved in the active center of a family of pyrophosphate-releasing NTPases. As a signature pattern the core region of

the domain was selected; it contains four conserved glutamate residues.

Consensus pattern: G-x(5)-E-x(4)-(STAGC)-(LIVMAC)-x-R-E-  
5 [LIVMFT]-x-E-E-

[1] Michaels M.L., Miller J.H. J. Bacteriol. 174:6321-  
6325(1992).

[2] Koonin E.V. Nucleic Acids Res. 21:4847-4847(1993).

10 [3] Mejean V., Salles C., Bullions M.J., Bessman M.J.,  
Claverys J.-P. Mol. Microbiol. 11:323-330(1994).

[4] Sakumi K., Furuichi M., Tsuzuki T., Kakuma T., Kawabata  
S., Maki H., Sekiguchi M. J. Biol. Chem. 268:23524-  
23530(1993).

15 [5] Thorne N.M.H., Hankin S., Wilkinson M.C., Nunez C.,  
Barracough R., McLennan A.G. Biochem. J. 311:717-721(1995).

#### 4. Cystatin domain

This is a very diverse family. Attempts to define separate  
20 subfamilies have failed. Typically, either the N-terminal or  
C-terminal end is very divergent. But splitting into two  
domains would make very short families. Cathelicidins are  
related to this family but have not been included. Number of  
members: 147

25 Inhibitors of cysteine proteases [1,2,3], which are  
found in the tissues and body fluids of animals, in the larva  
of the worm *Onchocerca volvulus* [4], as well as in plants,  
can be grouped into three distinct but related families:

- Type 1 cystatins (or stefins), molecules of about 100  
30 amino acid residues with neither disulfide bonds nor  
carbohydrate groups.
- Type 2 cystatins, molecules of about 115 amino acid  
residues which contain one or two disulfide loops near  
their C-terminus.

- Kininogens, which are multifunctional plasma  
glycoproteins.

They are the precursor of the active peptide bradykinin  
and play a role in blood coagulation by helping to position  
5 optimally prekallikrein and factor XI next to factor XII.

They are also inhibitors of cysteine proteases. Structurally,  
kininogens are made of three contiguous type-2 cystatin  
domains, followed by an additional domain (of variable  
length) which contains the sequence of bradykinin. The first  
10 of the three cystatin domains seems to have lost its  
inhibitory activity.

In all these inhibitors, there is a conserved region of  
five residues which has been proposed to be important for the  
binding to the cysteine proteases. The consensus pattern  
15 starts one residue before this conserved region.

-Consensus pattern: [GSTEQRV]-Q-[LIVT]-[VAF]-[SAGQ]-G-x-  
[LIVMVK]-x(2)-[LIVMFY]-x-[LIVMFYA]-[DENQXRHSIV]

20 [1] Barrett A.J. Trends Biochem. Sci. 12:193-196(1987).

[2] Rawlings N.D., Barrett A.J. J. Mol. Evol. 30:60-71(1990).

[3] Turk V., Bode W. FEBS Lett. 285:213-219(1991).

[4] Lustigman S., Brotman B., Huima T., Prince A.M. Mol.  
Biochem. Parasitol. 45:65-76(1991).

#### 5. Dehydrins signatures

A number of proteins are produced by plants that experience  
water-stress. Water-stress takes place when the water  
available to a plant falls below a critical level. The plant  
hormone abscisic acid (ABA) appears to modulate the response  
of plant to water-stress. Proteins that are expressed during  
water-stress are called dehydrins [1,2] or LEA group 2  
proteins [3]. The proteins that belong to this family are  
listed below.

59

- Arabidopsis thaliana XERO 1, XERO 2 (LTI30), RAB18, ERD10 (LTI45) ERD14 and COR47.
- Barley dehydrins B8, B9, B17, and B18.
- Cotton LEA protein D-11.
- Craterostigma plantagineum desiccation-related proteins A and B.
- Maize dehydrin M3 (RAB-17).
- Pea dehydrins DHN1, DHN2, and DHN3.
- Radish LEA protein.
- Rice proteins RAB 16B, 16C, 16D, RAB21, and RAB25.
- Tomato TAS14.
- Wheat dehydrin RAB 15 and cold-shock protein ccr4.0, cs66 and cs120.

5

Dehydrins share a number of structural features.

- One of the most notable features is the presence, in their central region, of a continuous run of five to nine serines followed by a cluster of charged residues. Such a region has been found in all known dehydrins so far with the exception of pea dehydrins. A second conserved feature is the presence of two copies of lysine-rich octapeptide; the first copy is located just after the cluster of charged residues that follows the poly-serine region and the second copy is found at the C-terminal extremity. Signature patterns for both regions were derived.

20

Consensus pattern: S(5)-[DE]-x-[DE]-G-x(1,2)-G-x(0,1)-[KR](4)  
Consensus pattern: [KR]-[LIM]-K-[DE]-K-[LIM]-P-G-

- [1] Close T.J., Kortt A.A., Chandler P.M. Plant Mol. Biol. 13:95-108(1989).
- [2] Robertson M., Chandler P.M. Plant Mol. Biol. 19:1031-1044(1992).

30

60

- [3] Dure L. III, Crouch M., Harada J., Ho T.-H. D., Mundy J., Quatrano R., Thomas T., Sung Z.R. Plant Mol. Biol. 12:475-486(1989).

- 5 6. D-isomer specific 2-hydroxyacid dehydrogenases (2-HADH)

This Pfam covers the Formate dehydrogenase, D-glycerate dehydrogenase and D-lactate dehydrogenase families in SCOP. A number of NAD-dependent 2-hydroxyacid dehydrogenases which seem to be specific for the D-isomer of their substrate have been shown [1,2,3,4] to be functionally and structurally related. These enzymes are listed below.

10

- D-lactate dehydrogenase (EC 1.1.1.28), a bacterial enzyme which catalyzes the reduction of D-lactate to pyruvate.

15

- D-glycerate dehydrogenase (EC 1.1.1.29) (NADH-dependent hydroxypyruvate reductase), a plant leaf peroxisomal enzyme that catalyzes the reduction of hydroxypyruvate to glycerate. This reaction is part of the glycolate pathway of photorespiration.

20

- D-glycerate dehydrogenase from the bacteria *Hyphomicrobium methylovorum* and *Methylobacterium extorquens*.

25

- 3-phosphoglycerate dehydrogenase (EC 1.1.1.95), a bacterial enzyme that catalyzes the oxidation of D-3-phosphoglycerate to 3-phosphohydroxypyruvate. This reaction is the first committed step in the 'phosphorylated' pathway of serine biosynthesis.

30

- Erythronate-4-phosphate dehydrogenase (EC 1.1.1.1-) (gene *pdxB*), a bacterial enzyme involved in the biosynthesis of pyridoxine (vitamin B6).
- D-2-hydroxyisocaproate dehydrogenase (EC 1.1.1.1-) (*D-hicDH*), a bacterial enzyme that catalyzes the reversible and stereospecific interconversion between

61

2-ketocarboxylic acids and D-2-hydroxy-carboxylic acids.

- Formate dehydrogenase (EC 1.2.1.2) (FDH) from the bacteria *Pseudomonas* sp. 101 and various fungi [5].
- Vancomycin resistance protein vanH from *Enterococcus faecium*; this protein is a D-specific alpha-keto acid dehydrogenase involved in the formation of a peptidoglycan which does not terminate by D-alanine thus preventing vancomycin binding.
- *Escherichia coli* hypothetical protein ycdW.
- *Escherichia coli* hypothetical protein ylaE.
- *Haemophilus influenzae* hypothetical protein H11556.
- Yeast hypothetical protein YER081w.
- Yeast hypothetical protein YIL074w.

All these enzymes have similar enzymatic activities and are structurally related. Three of the most conserved regions of these proteins have been selected to develop patterns. The first pattern is based on a glycine-rich region located in the central section of these enzymes; this region probably corresponds to the NAD-binding domain. The two other patterns contain a number of conserved charged residues, some of which may play a role in the catalytic mechanism.

-Consensus pattern: [LIWA]-[AG]-[IVT]-[LIVMY]-[AG]-x-G-  
25 [NHKRGSSAC]-[LIV]-G-x(13,14)-[LIVEMT]-x(2)-[FYWCTH]-[DNSTK]  
-Consensus pattern: [LIVMFYA]-[LIVFYWC]-x(2)-[SAC]-[DNQHR]-  
[IVFA]-[LIVF]-x-[LIVF]-[HNT]-x-?-x(4)-[STN]-x(2)-[LIVMF]-x-  
[GSDN]  
-Consensus pattern: [LMFATC]-[KPK]-x-[GSTDN]-x-[LIVMFYWR]-  
30 [LIVMFYW](2)-N-x-[STAGC]-R-[GP]-x-[LIVH]-[LIVMC]-[DNV]

[1] Grant G.A. Biochem. Biophys. Res. Commun. 165:1371-1374(1989).

62

- [2] Kochhar S., Hunziker P., Leong-Morgenthaler P.M., Hettlinger H. Biochem. Biophys. Res. Commun. 184:60-66(1992).
- [3] Ohta T., Taguchi H. J. Biol. Chem. 266:12588-12594(1991).
- [4] Goldberg J.D., Yoshida T., Brick P. J. Mol. Biol. 236:1123-1140(1994).

- [5] Popov V.O., Lamzin V.S. Biochem. J. 301:625-643(1994).

#### 7. dnaJ domains signatures and profile

The prokaryotic heat shock protein dnaJ interacts with the chaperone hsp70-like dnaK protein [1]. Structurally, the dnaJ protein consists of an N-terminal conserved domain (called 'J' domain) of about 70 amino acids, a glycine-rich region ('G' domain') of about 30 residues, a central domain containing four repeats of a CXXCXGKG motif ('CRR' domain) and a C-terminal region of 120 to 170 residues. Such a structure is shown in the following schematic representation:

```

+-----+-----+-----+-----+-----+-----+
terminal | +-----+ | N-terminal | | Gly-R | | CXXCXGKG | C-
+-----+-----+-----+-----+-----+-----+

```

It has been shown [2] that the 'J' domain as well as the 'CRR' domain are also found in other prokaryotic and eukaryotic proteins which are listed below.

#### a) Proteins containing both a 'J' and a 'CRR' domain:

- Yeast protein MASS/YDJ1 which seems to be involved in mitochondrial protein import.
- Yeast protein MDJ1, involved in mitochondrial biogenesis and protein folding.
- Yeast protein SCJ1, involved in protein sorting.
- Yeast protein XDJ1.
- Plants dnaJ homologs (from leek and cucumber).
- Human HDJ2, a dnaJ homolog of unknown function.
- Yeast hypothetical protein YNL077w.

#### b) Proteins containing a 'J' domain without a 'CRR' domain:

63

- *Rhizobium fredii* nolC, a protein involved in cultivar-specific nodulation of soybean.
- *Escherichia coli* cbpA [3], a protein that binds curved DNA.

5

- Yeast protein SEC63/NPL1, important for protein assembly into the endoplasmic reticulum and the nucleus.
- Yeast protein SISI, required for nuclear migration during mitosis.

10

- Yeast protein CAJ1.
- Yeast hypothetical protein YER041c.
- Yeast hypothetical protein YIR004w.
- Yeast hypothetical protein YJL162c.

15

- *Plasmodium falciparum* ring-infected erythrocyte surface antigen (RESA). RESA, whose function is not known, is associated with the membrane skeleton of newly invaded erythrocytes.

- Human HDJ1.

- Human HSF1, a neuronal protein.

20

- *Drosophila* cysteine-string protein (csp).

A signature pattern for the 'J' domain was developed, based on conserved positions in the C-terminal half of this domain. A pattern for the 'CRR' domain, based on the first two copies of that motif was also developed. A profile for the 'J' domain was also developed.

25

Consensus pattern: [FY]-x(2)-[LIVMA]-x(3)-[FYWHNT]-[DENQSA]-x-L-x-[DN]-x(3)-[KR]-x(2)-[FYI]-  
 Consensus pattern: C-[DEGSTHKR]-x-C-x-G-x-[GK]-[AGSDM]-x(2)-  
 [GSNKR]-x(4,6)-C-x(2,3)-C-x-G-x-G-

30

[1] Cyr D.M., Langer T., Douglas M.G. Trends Biochem. Sci. 19:176-181(1994).

64

[2] Bork P., Sander C., Valencia A., Bukau B. Trends Biochem. Sci. 17:129-129(1992).

[3] Ueguchi C., Kaneda M., Yamada H., Mizuno T. Proc. Natl. Acad. Sci. U.S.A. 91:1054-1058(1994).

5

8. Domain of unknown function

9. Gamma-thionins family signature

The following small plant proteins are evolutionary

10

related:

- Gamma-thionins from wheat endosperm (gamma-purothionins) and barley (gamma-hordothionins) which are toxic to animal cells and inhibit protein synthesis in cell free systems [1].

15

- A flower-specific thionin (FST) from tobacco [2].

- Antifungal proteins (AFP) from the seeds of Brassicaceae species such as radish, mustard, turnip and Arabidopsis thaliana [3].

- Inhibitors of insect alpha-amylases from sorghum [4].
- Probable protease inhibitor P322 from potato.

20

- A germination-related protein from cowpea [5].

- Anther-specific protein SF18 from sunflower [6]. SF18 is a protein that contains a gamma-thionin domain at its N-terminus and a proline-rich C-terminal domain.

25

- Soybean sulfur-rich protein SE60 [7].

- Vicia faba antibacterial peptides fabatin-1 and -2.

In their mature form, these proteins generally consist of about 45 to 50 amino-acid residues. As shown in the following schematic representation, these peptides contain eight conserved cysteines involved in disulfide bonds.

30

```

+-----+-----+-----+-----+-----+-----+
+-----+ | | | | |
xxCxxxxxxxxxCxxxxxxxxxCxxxxxxxxxCxxxxxxxxxCxxxxC
*****|*****|*****| | +---|-----+ +-----+
+-----+

```

35

65

'C': conserved cysteine involved in a disulfide bond.

':': position of the pattern.

Consensus pattern: [KRG]-x-C-x(3)-(SV)-x(2)-[FYWH]-x-[GF]-x-C-x(5)-C-x(3)-C [The four C's are involved in disulfide bonds]-

[1] Bruix M., Jimenez M.A., Santoro J., Gonzalez C., Collilla F.J., Mendez E., Rico M. *Biochemistry* 32:715-724(1993).

[2] Gu Q., Kawata E.E., Morse M.-J., Wu H.-M., Cheung A.Y. *Mol. Gen. Genet.* 234:89-96(1992).

[3] Terras F.R.G., Torreken S., van Leuven F., Osborn R.W., Vanderleyden J., Cammue B.P.A., Broekaert W.F. *FEBS Lett.* 316:233-240(1993).

[4] Bloch C. Jr., Richardson M. *FEBS Lett.* 279:101-104(1991).

[5] Ishibashi N., Yamauchi C., Minamikawa T. *Plant Mol. Biol.* 15:59-64(1990).

[7] Choi Y., Choi Y.D., Lee J.S. *Plant Physiol.* 101:699-700(1993).

#### 10. haloacid dehalogenase-like hydrolase

This family is structurally different from the alpha/beta hydrolase family (abhydrolase). This family includes L-2-haloacid dehalogenase, epoxide hydrolases and phosphatases.

The structure of the family consists of two domains. One is an inserted four helix bundle, which is the least well conserved region of the alignment, between residues 16 and 96 of Swiss:P24069. The rest of the fold is composed of the core alpha/beta domain.

[1] Hisano T, Hata Y, Fujii T, Liu JQ, Kurihara T, Esaki N, Soda K, *J Biol Chem* 1996; 271:20322-20330.

#### 11. Helix-turn-helix (HTH3)

66

This large family of DNA binding helix-turn helix proteins includes Cro Swiss:P03036 and CI Swiss:P03034.

12. Heme-binding domain in cytochrome b5 and oxidoreductases (heme\_1)

Cytochrome b5 is a membrane-bound hemo protein which acts as an electron carrier for several membrane-bound oxygenases [1]. There are two homologous forms of b5, one found in microsomes and one found in the outer membrane of mitochondria. Two conserved histidine residues serve as axial ligands for the heme group. The structure of a number of oxidoreductases consists of the juxtaposition of a heme-binding domain homologous to that of b5 and either a flavodehydrogenase or a molybdopterin domain. These enzymes are:

- Lactate dehydrogenase (EC 1.1.2.3) [2], an enzyme that consists of a flavodehydrogenase domain and a heme-binding domain called cytochrome b2.

- Nitrate reductase (EC 1.6.6.1), a key enzyme involved in the first step of nitrate assimilation in plants, fungi and bacteria [3,4]. Consists of a molybdopterin domain (see <PDOC09484>), a heme-binding domain called cytochrome b557, as well as a cytochrome reductase domain.

- Sulfite oxidase (EC 1.8.3.1) [5], which catalyzes the terminal reaction in the oxidative degradation of sulfur-containing amino acids. Also consists of a molybdopterin domain and a heme-binding domain.

This family of proteins also includes:

- TU-368, a *Drosophila* muscle protein: of unknown function [6].

- Fission yeast hypothetical protein SpAC1F12.10c.

- Yeast hypothetical protein YMR073c.

- Yeast hypothetical protein YMR272c.



A segment was used which includes the first of the two histidine heme ligands, as a signature pattern for the heme-binding domain of cytochrome b5 family.

5 Consensus pattern: [FY]-[LIVMK]-x(2)-H-P-[GA]-G [H is a heme axial ligand]-

[1] Ozols J. Biochim. Biophys. Acta 997:121-130 (1989).

[2] Guiard B. EMBO J. 4:3265-3272 (1985).

10 [3] Calza R., Huttner E., Vincentz M., Rouze P., Galangau F., Vaucheret H., Cherel I., Meyer C., Kronenberger J., Caboche M. Mol. Gen. Genet. 209:552-562 (1987).

[4] Crawford N.M., Smith M., Bollissimo D., Davis R.W. Proc. Natl. Acad. Sci. U.S.A. 85:5006-5010 (1988).

15 [5] Guiard B., Lederer F. Eur. J. Biochem. 100:441-453 (1979).

[6] Levin R.J., Boychuk P.L., Croniger C.M., Kazzaz J.A.,

Rozek C.E. Nucleic Acids Res. 17:6349-6367 (1989).

13. KH domain

20 KH motifs probably bind RNA directly. Auto antibodies to Nova, a KH domain protein, cause paraneoplastic opsoclonus ataxia.

[1] Burd CG, Dreyfuss G, Science 1994;265:615-621.

[2] Musco G, Stier G, Joseph C, Castiglione Morelli MA,

25 Nilges M, Gibson TJ, Pastore A, Cell 1996;85:237-245.

14. MAPEG family (aka: FLAP/GST2/LTC4S family signature)

The following mammalian proteins are evolutionary related [1]:

30 - Leukotriene C4 synthase (EC 2.5.1.37) (gene LTC4S), an enzyme that catalyzes the production of LTC4 from LTA4.

- Microsomal glutathione S-transferase II (EC 2.5.1.18) (GST-II) (gene GST2), an enzyme that can also produces LTC4 from LTA4.

5 - 5-lipoxygenase activating protein (gene FLAP), a protein that seems to be required for the activation of 5-lipoxygenase.

These are proteins of 150 to 160 residues that contain three transmembrane segments. As a signature pattern, a conserved region between the first and second transmembrane domains was selected.

10

Consensus pattern: G-x(3)-F-E-R-V-[FY]-x-A-[NQ]-x-N-C

[1] Jakobsson P.-J., Mancini J.A., Ford-Hutchinson A.W. J.

15 Biol. Chem. 271:22203-22210 (1996).

15. Pathogenesis-related protein Bet v I family signature

A number of plant proteins, which all seem to be involved in pathogen defense response, are structurally related [1,2,3]. These proteins are:

20

- Bet v I, the major pollen allergen from white birch.

Bet v I is the main cause of type I allergic reactions in Europe, North America and USSR.

- Aln g I, the major pollen allergen from alder.

25 - Api G I, the major allergen from celery.

- Car b I, the major pollen allergen from hornbeam.

- Cor a I, the major pollen allergen from hazel.

- Mal d I, the major pollen allergen from apple.

- Asparagus wound-induced protein AoPR1.

30 - Kidney bean pathogenesis-related proteins 1 and 2.

- Parsley pathogenesis-related proteins PRL-1 and PRL-3.

- Pea disease resistance response proteins p149, p176 and DRG49-C.

- Pea abscisic acid-responsive proteins ABR17 and ABR18.

- Potato pathogenesis-related proteins STH-2 and STH-21.
- Soybean stress-induced protein SAM22.

These proteins are thought to be intracellularly located. They contain from 155 to 160 amino acid residues. As a signature pattern, a conserved region located in the third quarter of these proteins has been selected

Consensus pattern: G-x(2)-[L-YMF]-x(4)-E-x(2)-[CSTAEN]-x(8,9)-[GND]-G-[GS]-[CS]-x(2)-K-x(4)-[FY]-

- 10 [1] Breiteneder H., Pottenburger K., Bito A., Valenta R., Kraft D., Rumpold H., Scheiner O., Breitenbach M. EMBO J. 8:1935-1938 (1989).
- [2] Crowell D., John M.E., Russell D., Amasino R.M. Plant Mol. Biol. 18:459-466 (1992).
- 15 [3] Warner S.A.J., Scott R., Draper J. Plant Mol. Biol. 19:555-561 (1992).

16. Photosystem I psaG / psaK (PSI PSAK) proteins signature  
Photosystem I (PSI) [1] is an integral membrane protein complex that uses light energy to mediate electron transfer from plastocyanin to ferredoxin. It is found in the chloroplasts of plants and cyanobacteria. PSI is composed of at least 14 different subunits, two of which PSI-G (gene psaG) and PSI-K (gene psaK) are small hydrophobic proteins of about 7 to 9 Kd and evolutionary related [2]. Both seem to contain two transmembrane regions. Cyanobacteria seem to encode only for PSI-K.

As a signature pattern, the best-conserved region was selected which seems to correspond to the second transmembrane region.

-Consensus pattern: [GT]-F-x-[LIVM]-x-[DEA]-x(2)-[GA]-x-[GTA]-[SA]-x-G-H-x-[LIVM]-[GA]

- [1] Golbeck J.H. Biochim. Biophys. Acta 895:167-204 (1987).
- [2] Kjaerulff S., Andersen B., Nielsen V.S., Møller B.L., Okkels J.S. J. Biol. Chem. 268:18912-18916 (1993).

# 5 17. Plant lipid transfer protein family signature (LTP)

Plant cells contain proteins, called lipid transfer proteins (LTP) [1,2,3], which are able to facilitate the transfer of phospholipids and other lipids across membranes. These proteins, whose subcellular location is not yet known, could play a major role in membrane biogenesis by conveying phospholipids such as waxes or cutin from their site of biosynthesis to membranes unable to form these lipids. Plant LTP's are proteins of about 9 Kd (90 amino acids) which contain eight conserved cysteine residues all involved in disulfide bridges, as shown in the following schematic representation.

```

+-----+ +-----+ | | | |
*****

```

```

xCxxxxCxxxxCxxxxCxxxxCxxxxCxxxxCxxx | | | +----
-----+ +-----+

```

'C': conserved cysteine involved in a disulfide bond.

'\*': position of the pattern.

Consensus pattern: [LIVM]-[PA]-x(2)-C-x-[LIVM]-x-[LIVM]-x-[LIVMFY]-x-[LIVM]-[ST]-x(3)-[DN]-C-x(2)-[LIVM] [The two C's are involved in disulfide bonds]

- [1] Wirtz K.W.A. Annu. Rev. Biochem. 60:73-99 (1991).
- [2] Arondel V., Kader J.C. Experientia 46:579-585 (1990).
- 30 [3] Ohlrogge J.B., Browse J., Somerville C.R. Biochim. Biophys. Acta 1082:1-26 (1991).

18. Ribosomal protein S7e signature

71

A number of eukaryotic ribosomal proteins can be grouped on the basis of sequence similarities [-]. One of these families consists of:

- Mammalian S7.
- Xenopus S8.
- Insect S7.

- Yeast probable ribosomal protein S7 (N2212).
- Fission yeast probable ribosomal protein S7 (SpAC18G6.13c).

10 These proteins have about 200 amino acids. A highly conserved stretch of 14 residues which is located in the central section and which is rich in charged residues was selected as a signature pattern.

15 Consensus pattern: [KR]-L-x-R-E-L-E-K-K-F-[SAP]-x-[KR]-H  
[1] Salazar C.E., Mills-Hamm D.M., Kumar V., Collins F.H.  
Nucleic Acids Res. 21:4147-4147(1993).

20 19. Ribosomal protein L34 signature

Ribosomal protein L34 is one of the proteins from the large subunit of the prokaryotic ribosome. It is a small basic protein of 44 to 51 amino-acid residues [1]. L34 belongs to a family of ribosomal proteins which, on the basis of sequence similarities, groups: - Eubacterial L34.

- Red algal chloroplast L34. - Cyanelle L34.

A conserved region that corresponds to the N-terminal half of L34 has been selected as a signature pattern.

30 -Consensus pattern: K-[RG]-T-[FYWL]-[EQS]-x(5)-[KRHS]-x(4,5)-G-F-x(2)-R

- [ 1] Old I.G., Margarita D., Saint Girons I.  
Nucleic Acids Res. 20:6097-6097(1992).

72

20. Ribosomal protein L6 signatures

Ribosomal protein L6 is one of the proteins from the large ribosomal subunit. In *Escherichia coli*, L6 is known to bind directly to the 23S rRNA and is located at the aminoacyl-tRNA binding site of the peptidyltransferase center. It belongs to a family of ribosomal proteins which, on the basis of sequence similarities [1,2,3,4], groups: - Eubacterial L6.

- Algal chloroplast L6.
- Cyanelle L6.
- Archaeobacterial L6.
- Marchantia polymorpha mitochondrial L6.
- Yeast mitochondrial Yml6 (gene MRPL6).
- Mammalian L9.
- Drosophila L9.
- Plants L9.
- Yeast L9 (YL11).

While all the above proteins are evolutionary related it is very difficult to derive a pattern that will find them all. Two patterns were therefore created, the first to detect eubacterial, cyanelle and mitochondrial L6, the second to detect archaeobacterial L6 as well as eukaryotic L9.

-Consensus pattern: [PS]-[DENS]-x-Y-K-[GA]-K-G-[LIVM]

-Consensus pattern: Q-x(3)-[LIVM]-x(2)-[KR]-x(2)-R-x-F-x-D-G-[LIVM]-Y-[LIVM]-x(2)-[KR]

- [1] Suzuki K., Olvera J., Wool I.G. Gene 93:297-300(1990).

[2] Schwank S., Harrer R., Schueller H.-J., Schweizer E.  
Curr. Genet. 24:136-140(1993).

[3] Golden B.L., Ramakrishnan V., White S.W. EMBO J. 12:4901-4908(1993).

- [ 4] Otake E., Hashimoto T., Mizuta K., Suzuki K. Protein Seq. Data Anal. 5:301-313(1993).

73

21. Ribosomal protein S14p/S29e (Ribosomal protein S14 signature)

Ribosomal protein S14 is one of the proteins from the small ribosomal subunit. In *Escherichia coli*, S14 is known to be required for the assembly of 30S particles and may also be responsible for determining the conformation of 16S rRNA at the A site. It belongs to a family of ribosomal proteins which, on the basis of sequence similarities [1,2], groups:

- Eubacterial S14.

- 10 - Algal and plant chloroplast S14.

- Cyanelle S14.

- Archaeobacterial *Methanococcus vannielii* S14.

- Plant mitochondrial S14.

- Yeast mitochondrial MRP2.

- Mammalian S29.

- Yeast YS29A/B.

S14 is a protein of 53 to 115 amino-acid residues. Our signature pattern is based on the few conserved positions located in the center of these proteins.

20

Consensus pattern: [RP]-x(0,1)-C-x(11,12)-[LIVMF]-x-[LIVMF]-[SC]-[RG]-x(3)-[RN]

- 25 [1] Chan Y.-L., Suzuki K., Olvera J., Wool I.C. Nucleic Acids Res. 21:649-655(1993).

[2] Otaka E., Hashimoto T., Mizuta K. Protein Seq. Data Anal. 5:285-300(1993).

22. Ribosomal protein S16 signature

30 Ribosomal protein S16 is one of the proteins from the small ribosomal subunit. It belongs to a family of ribosomal proteins which, on the basis of sequence similarities [1], groups:

- Eubacterial S16.

74

- Algal and plant chloroplast S16.

- Cyanelle S16.

- *Neurospora crassa* mitochondrial S24 (cyt-21).

5 S16 is a protein of about 100 amino-acid residues. A conserved region located in the N-terminal extremity of these proteins has been selected as a signature pattern.

Consensus pattern: [LIVMT]-x-[LIVM]-[KR]-L-[STAK]-R-x-G-[AKR]

- 10 [1] Otaka E., Hashimoto T., Mizuta K. Protein Seq. Data Anal. 5:285-300(1993).

23. Ribosomal protein S21 signature

Ribosomal protein S21 is one of the proteins from the small ribosomal subunit. So far S21 has only been found in eubacteria. It is a protein of 55 to 70 amino-acid residues. A conserved region in the N-terminal section of the protein has been selected as a signature pattern.

- 20 Consensus pattern: [DF]-x-A-[LIY]-[KR]-R-F-K-[KR]-x(3)-[KR]

24. Universal stress protein family (Usp)

By a wide range of stress conditions members of the Usp family are predicted to be related to the MADS-box proteins transcript fact and bind to DNA [2]. Number of members: 39

[1] Expression and role of the universal stress protein, UspA, of *Escherichia coli* during growth arrest. Nystrom T, Neidhardt FC; Mol Microbiol 1994; 11:537-544.

[2] Sequence analysis of eukaryotic developmental proteins: ancient and novel domains. Mushegian AR, Koonin EV; Genetics 1996; 144:817-828.

### III. Methods of Modulating Polypeptide Production

Within the scope of invention are chimeric gene constructs wherein the promoter and the structural coding sequence and/or other regulatory sequences within said constructs are heterologous to each other. "Heterologous sequences" are those that are not operatively linked or are not contiguous to each other in nature. For example, a promoter from corn is considered heterologous to an *Arabidopsis* coding region sequence. Also, a promoter from a gene encoding a growth factor from corn is considered heterologous to a sequence encoding the corn receptor for the growth factor. Regulatory element sequences, such as UTRs or 3' end termination sequences that do not originate in nature from the same gene as the coding sequence originates from, are considered heterologous to said coding sequence. On the other hand, elements operatively linked in nature are not heterologous. Thus, the promoter and coding portion of a corn gene expressing an amino acid transporter are not heterologous to each other.

Such chimeric polynucleotides are of particular interest for modulating gene expression in a host cell upon transformation of said cell with said chimeric polynucleotide.

Also within the scope of the invention are DNA molecules, whereof at least a part or portion of these DNA molecules are presented in SEQ TABLES 1 AND 2 of the present application, and wherein the structural coding sequence is under the control of its own promoter and/or its own regulatory elements. Such DNA molecules are useful for transforming the genome of a host cell or an organism regenerated from said host cell.

Typically, such polynucleotides, whether chimeric or not, are "exogenous to" the genome of an individual host cell or the organism regenerated from said host cell, such as a plant cell, respectively for a plant, when initially or subsequently introduced into said host cell or organism, by any means other

than by a sexual cross. Examples of means by which this can be accomplished are described below, and include *Agrobacterium*-mediated transformation (of dicots - e.g. Salomon et al. *EMBO J.* 3:141 (1984); Herrera-Estrella et al. *EMBO J.* 2:987 (1983); A.C. Vergunst et al. *Nucleic Acids Res.* 26:11, 2729 (1998); of monocots, representative papers are those by Escudero et al., *Plant J.* 10:355 (1996), Ishida et al., *Nature Biotechnology* 14:745 (1996), May et al., *Bio/Technology* 13:486 (1995)), biolistic methods (Armaleo et al., *Current Genetics* 17:97 (1990)), electroporation, in planta techniques, and the like. Such a plant containing the exogenous nucleic acid is referred to here as an  $R_1$  generation transgenic plant. Transgenic plants which arise from a sexual cross with another parent line or by selfing are "descendants or the progeny" of a  $R_1$  plant and are generally called  $F_n$  plants or  $S_n$  plants, respectively, meaning the number of generations.

The SDFs prepared as described herein can be used to prepare expression cassettes useful in a number of techniques for suppressing or enhancing expression.

#### III.A. Suppression

Expression cassettes of the invention can be used to suppress expression of endogenous genes which comprise the SDF sequence. Inhibiting expression can be useful, for instance, to tailor the ripening characteristics of a fruit (Oeller et al., *Science* 254:437 (1991)) or to influence seed size (WO98/07842) or to provoke cell ablation (Mariani et al., *Nature* 357: 384-387 (1992)).

As described above, a number of methods can be used to inhibit gene expression in plants, such as antisense, ribozyme, introduction of "exogenous" genes into a host cell, insertion of a polynucleotide sequence into the coding sequence and/or the promoter of the endogenous gene of interest, and the like.

### III.A.1. Antisense

An expression cassette as described above can be transformed into host cell or plant to produce an antisense strand of RNA. In plant cells, it has been suggested that antisense RNA inhibits gene expression by preventing the accumulation of mRNA which encodes the enzyme of interest, see, e.g., Sheehy et al., *Proc. Nat. Acad. Sci. USA*, 85:8805 (1988), and Hiatt et al., U.S. Patent No. 4,801,340.

### III.A.2. Ribozymes

Similarly, ribozyme constructs can be transformed into a plant to cleave mRNA and down-regulate translation.

### III.A.3. Co-Suppression

Another method of suppression is by introducing an exogenous copy of the gene to be suppressed. Introduction of expression cassettes in which a nucleic acid is configured in the sense orientation with respect to the promoter has been shown to be an effective means by which to block the transcription of target genes. A detailed description of this method is described above.

III.A.4. Insertion of Sequences into the Gene to be Modulated

Yet another means of suppressing gene expression is to insert a polynucleotide into the gene of interest to disrupt transcription or translation of the gene.

Homologous recombination could be used to target a polynucleotide insert to a gene using the Cre-Lox system (A.C. Vergunst et al., *Nucleic Acids Res.* 26:2729 (1998), A.C. Vergunst et al., *Plant Mol. Biol.* 38:393 (1998), H. Albert et al., *Plant J.* 7:649 (1995)).

In addition, random insertion of polynucleotides into a host cell genome can also be used to disrupt the gene of interest. Azpiroz-Leehan et al., *Trends in Genetics* 13:152 (1997). In this method, screening for clones from a library containing random insertions is preferred for identifying those that have polynucleotides inserted into the gene of interest. Such screening can be performed using probes and/or primers described above based on sequences from SEQ TABLES 1 AND 2, fragments thereof, and substantially similar sequence thereto. The screening can also be performed by selecting clones or  $R_1$  plants having a desired phenotype.

### III.A.5. Promoter Modulation

Inactivation of the promoter that drives a gene of interest can modulate transcription and translation, and therefore expression. For example, triple helices can be formed using oligonucleotides based on sequences from SEQ TABLES 1 AND 2, fragments thereof, and substantially similar sequence thereto. The oligonucleotide can be delivered to the host cell can bind to the promoter in the genome to form a triple helix and prevent transcription.

Additionally, a vector capable of producing the oligonucleotide can be inserted into the host cell to deliver the oligonucleotide.

### III.A.6. Expression of Mutants

An alternative method for inhibiting gene function is through the use of dominant negative mutations. Dominant negative mutations produce a mutant polypeptide which is capable of competing with the native polypeptide, but which does not produce the native result. Consequently, over expression of these mutations can titrate out an undesired activity of the native protein. For example, the inactive dominant-negative mutant may bind to the same receptor as the

79

native protein, preventing the native protein from activating a signal transduction pathway. Alternatively, the dominant-negative mutant can be an inactive enzyme still capable of binding to the same substrate as the native protein.

5 Dominant-negative mutants also can act upon the native protein itself to prevent activity. For example, the native protein may be active only as a homo-multimer or as one subunit of a hetero-multimer. Incorporation of an inactive subunit into the multimer with native subunit(s) can inhibit activity.

Thus, gene function can be modulated by insertion of an expression construct encoding a dominant-negative mutant into a host cell of interest.

### III.B. Enhanced Expression

15 Enhanced expression of a gene of interest in a host cell can be accomplished by either (1) insertion of an exogenous gene; or (2) promoter modulation.

#### III.B.1. Insertion of an Exogenous Gene

20 Insertion of an expression construct encoding an exogenous gene can boost the number of gene copies expressed in a host cell.

Such expression constructs can comprise genes that either encode the native protein that is of interest or that encode a variant that exhibits enhanced activity as compared to the native protein. Such genes encoding proteins of interest can be constructed from the sequences from SEQ TABLES 1 AND 2, fragments thereof, and substantially similar sequence thereto.

30 Such an exogenous gene can include either a constitutive promoter permitting expression in any cell in a host organism or a promoter that directs expression only in particular cells or times during a host cell life cycle or in response to environmental stimuli.

80

### III.3.2. Promoter Modulation

Some promoters require binding of a regulatory protein to be activated. Other promoters may need a protein that signals a promoter binding protein to expose a polymerase binding site. In either case, over-expression of such proteins can be used to enhance expression of a gene of interest by increasing the activation time of the promoter.

10 Such regulatory proteins are encoded by some of the sequences in SEQ TABLES 1 AND 2, fragments thereof, and substantially similar sequences thereto.

Coding sequences for these proteins can be constructed as described above.

In some cases, duplication of enhancer elements or insertion of exogenous enhancer elements will increase expression of a desired gene from a particular promoter. The useful enhancer elements can be portions of one or more of the SDFs of SEQ TABLES 1 AND 2.

### IV. Gene Constructs and Vector Construction

15 To use isolated SDFs of the present invention or a combination of them or parts and/or mutants and/or fusions of said SDFs in the above techniques, recombinant DNA vectors which comprise said SDFs and are suitable for transformation of cells, such as plant cells, are usually prepared.

20 The vector backbone can be any of those typical in the art such as plasmids, viruses, artificial chromosomes, BACs, YACs and PACs and vectors of the sort described by \*\*. Typically, a vector will comprise the exogenous gene, which in its turn comprises an SDF of the present invention to

25 be introduced into the genome of a host cell, and which gene may be an antisense construct, a ribozyme construct, or a structural coding sequence with any desired transcriptional and/or translational regulatory sequences, such as promoters

81

and 3' end termination sequences. Vectors of the invention can also include origins of replication, markers, homologous sequences, introns, etc.

A DNA sequence coding for the desired polypeptide, for example a cDNA sequence encoding a full length protein, will preferably be combined with transcriptional and translational initiation regulatory sequences which will direct the transcription of the sequence from the gene in the intended tissues of the transformed plant.

For example, for over-expression, a plant promoter fragment may be employed that will direct expression of the gene in all tissues of a regenerated plant. Such promoters are referred to herein as "constitutive" promoters and are active under most environmental conditions and states of development or cell differentiation. Examples of constitutive promoters include the cauliflower mosaic virus (CaMV) 35S transcription initiation region, the 1' or 2' promoter derived from T-DNA of *Agrobacterium tumefaciens*, and other transcription initiation regions from various plant genes known to those of skill.

Alternatively, the plant promoter may direct expression of an SDF of the invention in a specific tissue (tissue-specific promoters) or may be otherwise under more precise environmental control (inducible promoters). Examples of tissue-specific promoters under developmental control include promoters that initiate transcription only in certain tissues, such as root, ovule, fruit, seeds, or flowers. The promoter from a *LEC1* gene, described in copending application U.S. Ser. No. 09/103,478, is particularly useful for directing gene expression so that a desired gene product is located in embryos or seeds. Other suitable promoters include those from genes encoding storage proteins or the lipid body membrane protein, oleosin. A few root-specific promoters are noted above. Examples of environmental conditions that may affect

82

transcription by inducible promoters include anaerobic conditions, elevated temperature, or the presence of light.

If proper polypeptide expression is desired, a polyadenylation region at the 3'-end of the coding region should be included. The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from T-DNA.

The vector comprising the sequences (e.g., promoters or coding regions) from genes of the invention will typically comprise a marker gene that confers a selectable phenotype on plant cells. For example, the marker may encode biccide resistance, particularly antibiotic resistance, such as resistance to kanamycin, G418, bleomycin, hygromycin, or herbicide resistance, such as resistance to chlorosulfuron or phosphinotricin.

#### IV.A. Coding Sequences

Generally, the sequence in the transformation vector and to be introduced into the genome of the host cell does not need to be absolutely identical to an SDF of the present invention. Also, it is not necessary for it to be full length, relative to either the primary transcription product or fully processed mRNA. Use of sequences shorter than full-length may be preferred to avoid concurrent production of some plants that are overexpressors. Furthermore, the introduced sequence need not have the same intron or exon pattern as a native gene. Also, heterologous non-coding segments can be incorporated into the coding sequence without changing the desired amino acid sequence of the polypeptide to be produced.

#### IV.B. Promoters

As explained above, introducing an exogenous SDF from the same species or an orthologous SDF from another species can modulate the expression of a native gene corresponding to



that SDF of interest. Such an SDF construct can be under the control of either a constitutive promoter (e.g., the promoter of the 35S gene of the cauliflower mosaic virus or the promoter of the gene encoding the cowpea trypsin inhibitor) or a highly regulated inducible promoter (e.g., a copper inducible promoter). The promoter of interest can initially be either endogenous or heterologous to the species in question. When re-introduced into the genome of said species, such promoter becomes "exogenous" to said species. The promoter-SDF construct can be made using standard recombinant DNA techniques (Sambrook et al. 1989) and can be introduced to the species of interest by *Agrobacterium*-mediated transformation or by other means of transformation (e.g., particle gun bombardment) as referenced above. Over-expression of an SDF transgene can lead to co-suppression of the homologous gene thereby creating some alterations in the phenotypes of the transformed species as demonstrated by similar analysis of the chalcone synthase gene (Napoli et al., *Plant Cell* 2:279 (1990) and van der Krol et al., *Plant Cell* 2:291 (1990)). If an SDF is found to encode a protein with desirable characteristics, its over-expression can be controlled so that its accumulation can be manipulated in an organ- or tissue-specific manner utilizing a promoter having such specificity.

Likewise, if the promoter of an SDF (or an SDF that includes a promoter) is found to be tissue-specific or developmentally regulated, such a promoter can be utilized to drive the expression of a specific gene of interest (e.g., seed storage protein or root-specific protein). Thus, the level of accumulation of a particular protein can be manipulated or its spatial localization in an organ- or tissue-specific manner can be altered.

#### IV. C Signal Peptides

In some cases it may be desirable for the protein encoded by an introduced exogenous or orthologous SDF to be targeted (1) to a particular organelle, (2) to interact with a particular molecule or (3) for secretion outside of the cell harboring the introduced SDF. This will be accomplished using a signal peptide.

Signal peptides direct protein targeting, are involved in ligand-receptor interactions and act in cell to cell communication. Many proteins, especially soluble proteins, contain a signal peptide that targets the protein to one of several different intracellular compartments. In plants, these compartments include, but are not limited to, the endoplasmic reticulum (ER), mitochondria, plastids (such as chloroplasts), the vacuole, the Golgi apparatus, protein storage vesicles (PSV) and, in general, membranes. Some signal peptide sequences are conserved, such as the Asn-Pro-Ile-Arg amino acid motif found in the N-terminal propeptide signal that targets proteins to the vacuole (Marty (1999) *The Plant Cell* 11: 587-599). Other signal peptides do not have a consensus sequence *per se*, but are largely composed of hydrophobic amino acids, such as those signal peptides targeting proteins to the ER (Vitale and Denecke (1999) *The Plant Cell* 11: 615-628). Still others do not appear to contain either a consensus sequence or an identified common secondary sequence, for instance the chloroplast stromal targeting signal peptides (Keegstra and Cline (1999) *The Plant Cell* 11: 557-570). Furthermore, some targeting peptides are bipartite, directing proteins first to an organelle and then to a membrane within the organelle (e.g. within the thylakoid lumen of the chloroplast; see Keegstra and Cline (1999) *The Plant Cell* 11: 557-570). In addition to the diversity in sequence and secondary structure, placement of the signal peptide is also varied. Proteins destined for

the vacuole, for example, have targeting signal peptides found at the N-terminus, at the C-terminus and at a surface location in mature, folded proteins.

Signal peptides also serve as ligands for some receptors. Perhaps the best known example of this is the interaction of the ER targeting signal peptide with the signal recognition particle (SRP). Here, the SRP binds to the signal peptide, halting translation, and the resulting SRP complex then binds to docking proteins located on the surface of the ER, prompting the transfer of the protein into the ER.

These characteristics of signal proteins can be used to more tightly control the expression of introduced SDFs. In particular, associating the appropriate signal sequence with a specific SDF can allow sequestering of the protein in specific organelles (plastids, as an example), secretion outside of the cell, targeting interaction with particular receptors, etc. Hence, the inclusion of signal proteins in constructs involving the SDFs of the invention increases the range of manipulation of SDF expression. To carry this out, constructs are made with the nucleotide sequence of a known signal peptide immediately 5' to the initiation of the coding region of an SDF so that the signal peptide is translated in frame with the coding region and immediately precedes it.

The nucleotide sequence of the signal peptide can be isolated from characterized genes using common molecular biological techniques or can be synthesized in vitro.

#### V. Transformation Techniques

A wide range of techniques for inserting exogenous polynucleotides are known for a number of host cells, including, without limitation, bacterial, yeast, mammalian, insect and plant cells.

Techniques for transforming a wide variety of higher plant species are well known and described in the technical and scientific literature. See, e.g. Weising et al., *Ann. Rev. Genet.* 22:421 (1988); and Christou, *Euphytica*, v. 85, n.1-3:13-27, (1995).

DNA constructs of the invention may be introduced into the genome of the desired plant host by a variety of conventional techniques. For example, the DNA construct may be introduced directly into the genomic DNA of the plant cell using techniques such as electroporation and microinjection of plant cell protoplasts, or the DNA constructs can be introduced directly to plant tissue using ballistic methods, such as DNA particle bombardment. Alternatively, the DNA constructs may be combined with suitable T-DNA flanking regions and introduced into a conventional *Agrobacterium tumefaciens* host vector. The virulence functions of the *Agrobacterium tumefaciens* host will direct the insertion of the construct and adjacent marker into the plant cell DNA when the cell is infected by the bacteria (Vergunst et al., *Nucl. Acids. Res.* 26:2729 (1998); (site-directed integration using a Cre-Lox recombinase system); McCormac et al., *Mol. Biotechnol.* 8:199 (1997); Hamilton, *Gene* 200:107 (1997)); Salomon et al. *EMBO J.* 3:141 (1984); Herrera-Estrella et al. *EMBO J.* 2:987 (1983).

Microinjection techniques are known in the art and well described in the scientific and patent literature. The introduction of DNA constructs using polyethylene glycol precipitation is described in Paszkowski et al. *EMBO J.* 3:2717 (1984). Electroporation techniques are described in Fromm et al. *Proc. Natl. Acad. Sci. USA* 82:5824 (1985). Ballistic transformation techniques are described in Klein et al. *Nature* 327:773 (1987). *Agrobacterium tumefaciens*-mediated transformation techniques, including disarming and use of binary vectors, are well described in the scientific literature. See, for example Hamilton, *CM., Gene* 200:107

87

(1997); Muller et al. *Mol. Gen. Genet.* 207:171 (1987); Komari et al. *Plant J.* 10:165 (1996); Venkateswarlu et al. *Biotechnology* 9:1103 (1991) and Gleave, AP., *Plant Mol. Biol.* 20:1203 (1992); Graves and Goldman, *Plant Mol. Biol.* 7:34 (1986) and Gould et al., *Plant Physiology* 95:426 (1991).

Transformed plant cells which are derived by any of the above transformation techniques can be cultured to regenerate a whole plant that possesses the transformed genotype and thus the desired phenotype such as seedlessness. Such regeneration techniques rely on manipulation of certain phytohormones in a tissue culture growth medium, typically relying on a biocide and/or herbicide marker which has been introduced together with the desired nucleotide sequences. Plant regeneration from cultured protoplasts is described in Evans et al., *Protoplasts Isolation and Culture* in "Handbook of Plant Cell Culture," pp. 124-176, MacMillan Publishing Company, New York, 1983; and Binding, *Regeneration of Plants, Plant Protoplasts*, pp. 21-73, CRC Press, Boca Raton, 1988. Regeneration can also be obtained from plant callus, explants, organs, or parts thereof. Such regeneration techniques are described generally in Klee et al. *Ann. Rev. of Plant Phys.* 38:467 (1987). Regeneration of monocots (rice) is described by Hosoyama et al. (*Biosci. Biotechnol. Biochem.* 58:1500 (1994)) and by Ghosh et al. (*J. Biotechnol.* 32:1 (1994)). The nucleic acids of the invention can be used to confer desired traits on essentially any plant.

Thus, the invention has use over a broad range of plants, including species from the genera *Asparagus*, *Atropa*, *Avena*, *Brassica*, *Citrus*, *Citrullus*, *Capsicum*, *Cucumis*, *Cucurbita*, *Daucus*, *Fragaria*, *Glycine*, *Gossypium*, *Helianthus*, *Heterocallis*, *Hordeum*, *Hyoscyamus*, *Lactuca*, *Linum*, *Lolium*, *Lycopersicon*, *Malus*, *Manihot*, *Majorana*, *Medicago*, *Nicotiana*, *Oryza*, *Panicum*, *Pennisetum*, *Persea*, *Pisum*, *Pyrus*, *Prunus*, *Raphanus*, *Secale*, *Senecio*, *Sinapis*, *Solanum*, *Sorghum*, *Trigonella*, *Triticum*, *Vitis*, *Vigna*, and, *Zea*.

88

One of skill will recognize that after the expression cassette is stably incorporated in transgenic plants and confirmed to be operable, it can be introduced into other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed.

The particular sequences of SDFs identified are provided in the attached SEQ TABLES 1 AND 2. One of ordinary skill in the art, having this data, can obtain cloned DNA fragments, synthetic DNA fragments or polypeptides constituting desired sequences by recombinant methodology known in the art or described herein.

#### EXAMPLES

The invention is illustrated by way of the following examples. The invention is not limited by these examples as the scope of the invention is defined solely by the claims following.

#### EXAMPLE 1: SOUTHERN HYBRIDIZATIONS

The SDFs of the invention can be used in Southern hybridizations as described above. The following describes extraction of DNA from nuclei of plant cells, digestion of the nuclear DNA and separation by length, transfer of the separated fragments to membranes, preparation of probes for hybridization, hybridization and detection of the hybridized probe.

The procedures described herein can be used to isolate related polynucleotides or for diagnostic purposes. Moderate stringency hybridization conditions, as defined above, are described in the present example. These conditions result in detection of hybridization between sequences having at least 70% sequence identity. As described above, the hybridization and wash conditions can be changed to reflect the desired

89

degree of sequence identity between probe and target sequences that can be detected.

In the following procedure, a probe for the hybridization is produced from two PCR reactions using two primers from genomic sequence of *Arabidopsis thaliana*. As described above, the particular template for generating the probe can be any desired template.

The first PCR product is assessed to validate the size of the primer to assure it is of the expected size. Then the product of the first PCR is used as a template, with the same pair of primers used in the first PCR, in a second PCR that produces a labeled product used as the probe.

Fragments detected by hybridization, or other bands of interest, can be isolated from gels used to separate genomic DNA fragments by known methods for further purification and/or characterization.

#### Buffers for nuclear DNA extraction

##### 1. 10X HB

	1000 ml	
40 mM spermidine	10.2 g	Spermine (Sigma S-2876) and spermidine (Sigma S-2501)
10 mM spermine	3.5 g	Stabilize chromatin and the nuclear membrane
0.1 M EDTA (disodium)	37.2 g	EDTA inhibits nuclease
0.1 M Tris	12.1 g	Buffer
0.8 M KCl	59.6 g	Adjusts ionic strength for stability of nuclei

Adjust pH to 9.5 with 10 N NaOH. It appears that there is a nuclease present in leaves. Use of pH 9.5 appears to inactivate this nuclease.

90

##### 2. 2 M sucrose (684 g per 1000 ml)

Heat about half the final volume of water to about 50°C. Add the sucrose slowly then bring the mixture to close to final volume; stir constantly until it has dissolved. Bring the solution to volume.

##### 3. Sarkosyl solution (lyses nuclear membranes)

	1000 ml	
N-lauroyl sarcosine (Sarkosyl)	20.0 g	
0.1 M Tris	12.1 g	
0.04 M EDTA (Disodium)	14.9 g	

Adjust the pH to 9.5 after all the components are dissolved and bring up to the proper volume.

##### 4. 20% Triton X-100

80 ml Triton X-100

320 ml 1XHB (w/o  $\beta$ -ME and PMSF)

Prepare in advance; Triton takes some time to dissolve

##### A. Procedure

##### 1. Prepare 1X "H" buffer (keep ice-cold during use)

	1000 ml	
10X HB	100 ml	
2 M sucrose	250 ml	a non-ionic osmoticum
Water	634 ml	

Added just before use:

100 mM PMSF\*

10 ml a protease

inhibitor; protects

nuclear membrane proteins

91

$\beta$ -mercaptoethanol 1 ml inactivates nuclease by reducing disulfide bonds

\*100 mM PMSF

(phenyl methyl sulfonyl fluoride, Sigma P-7626)  
(add 0.0875 g to 5 ml 100% ethanol)

2. Homogenize the tissue in a blender (use 300-400 ml of 1xHB per blender). Be sure that you use 5-10 ml of HB buffer per gram of tissue. Blenders generate heat so be sure to keep the homogenate cold. It is necessary to put the blenders in ice periodically.

3. Add the 20% Triton X-100 (25 ml per liter of homogenate) and gently stir on ice for 20 min. This lyses plastid, but not nuclear, membranes.

4. Filter the tissue suspension through several nylon filters into an ice-cold beaker. The first filtration is through a 250-micron membrane; the second is through an 85-micron membrane; the third is through a 50-micron membrane; and the fourth is through a 20-micron membrane. Use a large funnel to hold the filters. Filtration can be sped up by gently squeezing the liquid through the filters.

5. Centrifuge the filtrate at 1200 x g for 20 min. at 4°C to pellet the nuclei.

6. Discard the dark green supernatant. The pellet will have several layers to it. One is starch; it is white and gritty. The nuclei are gray and soft. In the early

92

steps, there may be a dark green and somewhat viscous layer of chloroplasts.

Wash the pellets in about 25 ml cold H buffer (with Triton X-100) and resuspend by swirling gently and pipetting. After the pellets are resuspended.

Pellet the nuclei again at 1200 - 1300 x g. Discard the supernatant.

Repeat the wash 3-4 times until the supernatant has changed from a dark green to a pale green. This usually happens after 3 or 4 resuspensions. At this point, the pellet should be grayish white and very slippery. The Triton X-100 in these repeated steps helps to destroy the chloroplasts and mitochondria that contaminate the prep.

15 Resuspend the nuclei for a final time in a total of 15 ml of H buffer and transfer the suspension to a sterile 125 ml Erlenmeyer flask.

7. Add 15 ml, dropwise, cold 2% Sarkosyl, 0.1 M Tris, 0.04 M EDTA solution (pH 9.5) while swirling gently. This lyses the nuclei. The solution will become very viscous.

8. Add 30 grams of CsCl and gently swirl at room temperature until the CsCl is in solution. The mixture will be gray, white and viscous.

9. Centrifuge the solution at 11,400 x g at 4°C for at least 30 min. The longer this spin is, the firmer the protein pellicle.

10. The result should be a clear green supernatant over a white pellet, and (perhaps) under a protein pellicle. Carefully remove the solution under the protein pellicle and above the pellet. Determine the density of the solution by weighing 1 ml of solution and add CsCl if necessary to bring to 1.57 g/ml. The solution contains dissolved solids (sucrose etc) and the refractive index alone will not be an accurate guide to CsCl concentration.

11. Add 20  $\mu$ l of 10 mg/ml EtDr per ml of solution.

12. Centrifuge at 184,000  $\times$  g for 16 to 20 hours in a fixed-angle rotor.

13. Remove the dark red supernatant that is at the top of the tube with a plastic transfer pipette and discard. Carefully remove the DNA band with another transfer pipette. The DNA band should be visible in room light; otherwise, use a long wave UV light to locate the band.

14. Extract the ethidium bromide with isopropanol saturated with water and salt. Once the solution is clear, extract at least two more times to ensure that all of the EtBr is gone. Be very gentle, as it is very easy to shear the DNA at this step. This extraction may take a while because the DNA solution tends to be very viscous. If the solution is too viscous, dilute it with TE.

15. Dialyze the DNA for at least two days against several changes (at least three times) of TE (10 mM Tris, 1mM EDTA, pH 8) to remove the cesium chloride.

16. Remove the dialyzed DNA from the tubing. If the dialyzed DNA solution contains a lot of debris, centrifuge the DNA solution at least at 2500  $\times$  g for 10 min. and carefully transfer the clear supernatant to a new tube. Read the A260 concentration of the DNA.

17. Assess the quality of the DNA by agarose gel electrophoresis (1% agarose gel) of the DNA. Load 50 ng and 100 ng (based on the OD reading) and compare it with known and good quality DNA. Undigested lambda DNA and a lambda-HindIII-digested DNA are good molecular weight makers.

#### Protocol for Digestion of Genomic DNA

##### Protocol:

1. The relative amounts of DNA for different crop plants that provide approximately a balanced number of genome equivalent is given in Table 3. Note that due to the size of the wheat genome, wheat DNA will be underrepresented. Lambda DNA provides a useful control for complete digestion.

2. Precipitate the DNA by adding 3 volumes of 100% ethanol. Incubate at  $-20^{\circ}\text{C}$  for at least two hours. Yeast DNA can be purchased and made up at the necessary concentration, therefore no precipitation is necessary for yeast DNA.

3. Centrifuge the solution at 11,400  $\times$  g for 20 min. Decant the ethanol carefully (be careful not to disturb the pellet). Be sure that the residual ethanol is completely removed either by vacuum desiccation or by carefully wiping the sides of the tubes with a clean tissue.

95

4. Resuspend the pellet in an appropriate volume of water. Be sure the pellet is fully resuspended before proceeding to the next step. This may take about 30 min.

5. Add the appropriate volume of 10X reaction buffer provided by the manufacturer of the restriction-enzyme to the resuspended DNA followed by the appropriate volume of enzymes. Be sure to mix it properly by slowly swirling the tubes.

- 10 6. Set-up the lambda digestion-control for each DNA that you are digesting.

7. Incubate both the experimental and lambda digests overnight at 37°C. Spin down condensation in a microfuge before proceeding.

- 15 8. After digestion, add 2 µl of loading dye (typically 0.25% bromophenol blue, 0.25% xylene cyanol in 15% Ficoll or 30% glycerol) to the lambda-control digests and load in 1% TPE-agarose gel (TPE is 90 mM Tris-phosphate, 2 mM EDTA, pH 8). If the lambda DNA in the lambda control digests are completely digested, proceed with the precipitation of the genomic DNA in the digests.

9. Precipitate the digested DNA by adding 3 volumes of 100% ethanol and incubating in -20°C for at least 2 hours (preferably overnight).

EXCEPTION: Arabidopsis and yeast DNA are digested in an appropriate volume; they don't have to be precipitated.

96

10. Resuspend the DNA in an appropriate volume of TE (e.g., 22 µl x 50 blots = 1100 µl) and an appropriate volume of 10X loading dye (e.g., 2.4 µl x 50 blots = 120 µl). Be careful in pipetting the loading dye - it is viscous. Be sure you are pipetting the correct volume.

Table 3

Some guide points in digesting genomic DNA.

Species	Genome Size	Size Relative to Arabidopsis	Genome Equivalent to 2 µg Arabidopsis DNA	Amount of DNA per blot
Arabidopsis	120 Mb	1X	1X	2 µg
Brassica	1,100 Mb	9.2X	0.54X	10 µg
Corn	2,800 Mb	23.3X	0.43X	20 µg
Cotton	2,300 Mb	19.2X	0.52X	20 µg
Oat	11,300 Mb	94X	0.11X	20 µg
Rice	400 Mb	3.3X	0.75X	5 µg
Soybean	1,100 Mb	9.2X	0.54X	10 µg
Sugarbeet	758 Mb	6.3X	0.8X	10 µg
Sweetclover	1,100 Mb	9.2X	0.54X	10 µg
Wheat	16,000 Mb	133X	0.08X	20 µg
Yeast	15 Mb	0.12X	1X	0.25 µg

#### 10 Protocol for Southern Blot Analysis

The digested DNA samples are electrophoresed in 1% agarose gels in 1X TPE buffer. Low voltage; overnight separations are preferred. The gels are stained with EtBr and photographed.

1. For blotting the gels, first incubate the gel in 0.25 N HCl (with gentle shaking) for about 15 min.

2. Then briefly rinse with water. The DNA is denatured by 2 incubations. Incubate (with shaking) in 0.5 M NaOH in 1.5 M NaCl for 15 min.

3. The gel is then briefly rinsed in water and neutralized by incubating twice (with shaking) in 1.5 M Tris pH 7.5 in 1.5 M NaCl for 15 min.

4. A nylon membrane is prepared by soaking it in water for at least 5 min, then in 6X SSC for at least 15 min. before use. (20x SSC is 175.3 g NaCl, 88.2 g sodium citrate per liter, adjusted to pH 7.0.)

5. The nylon membrane is placed on top of the gel and all bubbles in between are removed. The DNA is blotted from the gel to the membrane using an absorbent medium, such as paper toweling and 6x SSC buffer. After the transfer, the membrane may be lightly brushed with a gloved hand to remove any agarose sticking to the surface.

6. The DNA is then fixed to the membrane by UV crosslinking and baking at 80°C. The membrane is stored at 4°C until use.

#### B. Protocol for PCR Amplification of Genomic Fragments in Arabidopsis

##### Amplification procedures:

- Mix the following in a 0.20 ml PCR tube or 96-well PCR plate:

Volume	Stock	Final Amount or Conc.
--------	-------	-----------------------

0.5 µl	~ 10 ng/µl genomic DNA <sup>1</sup>	5 ng
2.5 µl	10X PCR buffer	20 mM Tris, 50 mM KCl
0.75 µl	50 mM MgCl <sub>2</sub>	1.5 mM
1 µl	10 pmol/µl Primer 1 (Forward)	10 pmol
1 µl	10 pmol/µl Primer 2 (Reverse)	10 pmol
0.5 µl	5 mM dNTPs	0.1 mM
0.1 µl	5 units/µl Platinum Taq <sup>™</sup> (Life Technologies, Gaithersburg, MD) DNA Polymerase	1 units
(to 25 µl)	Water	

- The template DNA is amplified using a Perkin Elmer 9700 PCR machine:

- 94°C for 10 min. followed by

2)	3)	4)
5 cycles:	5 cycles:	25 cycles:
94°C - 30 sec	94°C - 30 sec	94°C - 30 sec
62°C - 30 sec	58°C - 30 sec	53°C - 30 sec
72°C - 3 min	72°C - 3 min	72°C - 3 min

- 72°C for 7 min. Then the reactions are stopped by chilling to 4°C.

<sup>1</sup> Arabidopsis DNA is used in the present experiment, but the procedure is a general one.



The procedure can be adapted to a multi-well format if necessary.

#### Quantification and Dilution of PCR Products:

1. The product of the PCR is analyzed by electrophoresis in a 1% agarose gel. A linearized plasmid DNA can be used as a quantification standard (usually at 50, 100, 200, and 400 ng). These will be used as references to approximate the amount of PCR products. HindIII-digested Lambda DNA is useful as a molecular weight marker. The gel can be run fairly quickly; e.g., at 100 volts. The standard gel is examined to determine that the size of the PCR products is consistent with the expected size and if there are significant extra bands or smeary products in the PCR reactions.

2. The amounts of PCR products can be estimated on the basis of the plasmid standard.

3. For the small number of reactions that produce extraneous bands, a small amount of DNA from bands with the correct size can be isolated by dipping a sterile 10- $\mu$ l tip into the band while viewing through a UV Transilluminator. The small amount of agarose gel (with the DNA fragment) is used in the labeling reaction.

#### C. Protocol for PCR-DIG-Labeling of DNA

##### Solutions:

- Reagents in PCR reactions (diluted PCR products, 10X PCR Buffer, 50 mM MgCl<sub>2</sub>, 5 U/ $\mu$ l Platinum Taq Polymerase, and the primers)

10X dNTP + DIG-11-dUTP (1:5): (2 mM dATP, 2 mM dCTP, 2 mM dGTP, 1.65 mM dTTP, 0.35 mM DIG-11-dUTP)

10X dNTP + DIG-11-dUTP (1:10): (2 mM dATP, 2 mM dCTP, 2 mM dGTP, 1.81 mM dTTP, 0.19 mM DIG-11-dUTP)

10X dNTP + DIG-11-dUTP (1:15): (2 mM dATP, 2 mM dCTP, 2 mM dGTP, 1.875 mM dTTP, 0.125 mM DIG-11-dUTP)

TE buffer (10 mM Tris, 1 mM EDTA, pH 8)

Maleate buffer: In 700 ml of deionized distilled water, dissolve 11.61 g maleic acid and 8.77 g NaCl. Add NaOH to adjust the pH to 7.5. Bring the volume to 1 L. Stir for 15 min. and sterilize.

10% blocking solution: In 80 ml deionized distilled water, dissolve 1.16g maleic acid. Next, add NaOH to adjust the pH to 7.5. Add 10 g of the blocking reagent powder (Boehringer Mannheim, Indianapolis, IN, Cat. no. 1096176). Heat to 60°C while stirring to dissolve the powder. Adjust the volume to 100 ml with water. Stir and sterilize.

1% blocking solution: Dilute the 10% stock to 1% using the maleate buffer.

Buffer 3 (100 mM Tris, 100 mM NaCl, 50 mM MgCl<sub>2</sub>, pH 9.5). Prepared from autoclaved solutions of 1M Tris pH 9.5, 5 M NaCl, and 1 M MgCl<sub>2</sub> in autoclaved distilled water.

Procedure:

1. PCR reactions are performed in 25 µl volumes containing:

- |                          |                                |
|--------------------------|--------------------------------|
| PCR buffer               | 1X                             |
| MgCl <sub>2</sub>        | 1.5 mM                         |
| 10X dNTP + DIG-11-dUTP   | 1X (please see the note below) |
| Platinum Taq™ Polymerase | 1 unit                         |
| 10 pg probe DNA          |                                |
| 10 pmol primer 1         |                                |

10 **Note:**

	Use for:
10X dNTP + DIG-11-dUTP (1:5)	< 1 kb
10X dNTP + DIG-11-dUTP (1:10)	1 kb to 1.8 kb
10X dNTP + DIG-11-dUTP (1:15)	> 1.8 kb

2. The PCR reaction uses the following amplification cycles:

- 1) 94°C for 10 min.

2)	3)	4)
5 cycles:	5 cycles:	25 cycles:
95°C - 30 sec	95°C - 30 sec	95°C - 30 sec
61°C - 1 min	59°C - 1 min	51°C - 1 min
73°C - 5 min	75°C - 5 min	73°C - 5 min

- 5) 72°C for 8 min. The reactions are terminated by chilling to 4°C (hold).

3. The products are analyzed by electrophoresis- in a 1% agarose gel, comparing to an aliquot of the unlabelled probe starting material.

4. The amount of DIG-labeled probe is determined as follows:

- 5 Make serial dilutions of the diluted control DNA in dilution buffer (TE: 10 mM Tris and 1 mM EDTA, pH 8) as shown in the following table:

DIG-labeled control DNA starting conc.	Stepwise Dilution	Final Conc. (Dilution Name)
5 ng/µl	1 µl in 49 µl TE	100 pg/µl (A)
100 pg/µl (A)	25 µl in 25 µl TE	50 pg/µl (B)
50 pg/µl (B)	25 µl in 25 µl TE	25 pg/µl (C)
25 pg/µl (C)	20 µl in 30 µl TE	10 pg/µl (D)

- a. Serial deletions of a DIG-labeled standard DNA ranging from 100 pg to 10 pg are spotted onto a positively charged nylon membrane, marking the membrane lightly with a pencil to identify each dilution.

- b. Serial dilutions (e.g., 1:50, 1:2500, 1:10,000) of the newly labeled DNA probe are spotted.

- c. The membrane is fixed by UV crosslinking.

- d. The membrane is wetted with a small amount of maleate buffer and then incubated in 1% blocking solution for 15 min at room temp.

- e. The labeled DNA is then detected using alkaline phosphatase conjugated anti-DIG antibody (Boehringer Mannheim, Indianapolis, IN, cat. no.

103

1093274) and an NBT substrate according to the manufacturer's instruction.

- f. Spot intensities of the control and experimental dilutions are then compared to estimate the concentration of the PCR-DIG-labeled probe.

#### D. Prehybridization and Hybridization of Southern Blots

##### Solutions:

100% Formamide purchased from Gibco

20X SSC (1X = 0.15 M NaCl, 0.015 M

Na<sub>2</sub>Citrate)

per L: 175 g NaCl

87.5 g Na<sub>2</sub>Citrate·2H<sub>2</sub>O

20% Sarkosyl (N-lauroyl-sarcosine)

20% SDS (sodium dodecyl sulphate)

- 15 10% Blocking Reagent: In 80 ml deionized distilled water, dissolve 1.16 g maleic acid. Next, add NaOH to adjust the pH to 7.5. Add 10 g of the blocking reagent powder. Heat to 60°C while stirring to dissolve the powder. Adjust the volume to 100 ml with water. Stir and sterilize.

##### Prehybridization Mix:

Final Concentration	Components	Volume (per 100 ml)	Stock
50%	Formamide	50 ml	100%
5X	SSC	25 ml	20X
0.1%	Sarkosyl	0.5 ml	20%

104

0.02%	SDS	0.1 ml	20%
2%	Blocking Reagent	20 ml	10%
	Water	4.4 ml	

##### General Procedures:

- Place the blot in a heat-sealable plastic bag and add an appropriate volume of prehybridization solution (30 ml/100cm<sup>2</sup>) at room temperature. Seal the bag with a heat sealer, avoiding bubbles as much as possible. Lay down the bags in a large plastic tray (one tray can accommodate at least 4-5 bags). Ensure that the bags are lying flat in the tray so that the prehybridization solution is evenly distributed throughout the bag. Incubate the blot for at least 2 hours with gentle agitation using a waver shaker.
- Denature DIG-labeled DNA probe by incubating for 10 min. at 98°C using the PCR machine and immediately cool it to 4°C.
- Add probe to prehybridization solution (25 ng/ml; 30 ml = 750 ng total probe) and mix well but avoid foaming. Bubbles may lead to background.
- Pour off the prehybridization solution from the hybridization bags and add new prehybridization and probe solution mixture to the bags containing the membrane.
- Incubate with gentle agitation for at least 16 hours.
- Proceed to medium stringency post-hybridization wash:

105

Three times for 20 min. each with gentle agitation using 1X SSC, 1% SDS at 60°C.

All wash solutions must be prewarmed to 60°C. Use about 100 ml of wash solution per membrane.

- 5 To avoid background keep the membranes fully submerged to avoid drying in spots; agitate sufficiently to avoid having membranes stick to one another.

7. After the wash, proceed to immunological detection and CSPD development.

#### 10 E. Procedure for Immunological Detection with CSPD

##### Solutions:

Buffer 1: Maleic acid buffer (0.1 M maleic acid, 0.15 M NaCl; adjusted to pH 7.5 with NaOH)

- 15 Washing buffer: Maleic acid buffer with 0.3% (v/v) Tween 20.

Blocking stock solution 10% blocking reagent in buffer 1.

- 20 Dissolve (10X concentration): blocking reagent powder (Boehringer Mannheim, Indianapolis, IN, cat. no. 1096176) by constantly stirring on a 65°C heating block or heat in a microwave, autoclave and store at 4°C.

- 25 Buffer 2

106

(1X blocking solution): Dilute the stock solution 1:10 in Buffer 1.

Detection buffer: 0.1 M Tris, 0.1 M NaCl, pH 9.5

##### Procedure:

- 5 1. After the post-hybridization wash the blots are briefly rinsed (1-5 min.) in the maleate washing buffer with gentle shaking.
2. Then the membranes are incubated for 30 min. in Buffer 2 with gentle shaking.

- 10 3. Anti-DIG-AP conjugate (Boehringer Mannheim, Indianapolis, IN, cat. no. 1093274) at 75 mU/ml (1:10,000) in Buffer 2 is used for detection. 75 ml of solution can be used for 3 blots.

- 15 4. The membrane is incubated for 30 min. in the antibody solution with gentle shaking.

5. The membrane are washed twice in washing buffer with gentle shaking. About 250 mls is used per wash for 3 blots.

- 20 6. The blots are equilibrated for 2-5 min in 60 ml detection buffer.

7. Dilute CSPD (1:200) in detection buffer. (This can be prepared ahead of time and stored in the dark at 4°C).

- 25 The following steps must be done individually. Bags (one for detection and one for exposure) should be cut and ready before doing the following steps.

107

8. The blot is carefully removed from the detection buffer and excess liquid removed without drying the membrane. The blot is immediately placed in a bag and 1.5 ml of CSPD solution is added. The CSPD solution can be spread over the membrane. Bubbles present at the edge and on the surface of the blot should be removed by gentle rubbing. The membrane is incubated for 5 min. in CSPD solution.

9. Excess liquid is removed and the membrane is blotted briefly (DNA side up) on Whatman 3MM paper. Do not let the membrane dry completely.

10. Seal the damp membrane in a hybridization bag and incubate for 10 min at 37°C to enhance the luminescent reaction.

11. Expose for 2 hours at room temperature to X-ray film. Multiple exposures can be taken. Luminescence continues for at least 24 hours and signal intensity increases during the first hours.

#### Example 2: Transformation of Carrot Cells

Transformation of plant cells can be accomplished by a number of methods, as described above. Similarly, a number of plant genera can be regenerated from tissue culture following transformation. Transformation and regeneration of carrot cells as described herein is illustrative.

Single cell suspension cultures of carrot (*Daucus carota*) cells are established from hypocotyls of cultivar Early Nantes in B<sub>3</sub> growth medium (O.L. Gamborg et al., *Plant Physiol.* 45:372 (1970)) plus 2,4-D and 15 mM CaCl<sub>2</sub> (B<sub>3</sub> -44 medium) by methods known in the art. The suspension cultures are subcultured by adding 10 ml of the suspension culture to

108

40 ml of B<sub>3</sub>-44 medium in 250 ml flasks every 7 days and are maintained in a shaker at 150 rpm at 27 °C in the dark.

The suspension culture cells are transformed with exogenous DNA as described by Z. Chen et al. *Plant Mol. Bio.* 36:163 (1998). Briefly, 4-days post-subculture cells are incubated with cell wall digestion solution containing 0.4 M sorbitol, 2% driselase, 5mM MES (2-[N-Morpholino] ethanesulfonic acid) pH 5.0 for 5 hours. The digested cells are pelleted gently at 60 xg for 5 min. and washed twice in W5 solution containing 154 mM NaCl, 5 mM KCl, 125 mM CaCl<sub>2</sub> and 5mM glucose, pH 6.0. The protoplasts are suspended in MC solution containing 5 mM MES, 20 mM CaCl<sub>2</sub>, 0.5 M mannitol, pH 5.7 and the protoplast density is adjusted to about 4 x 10<sup>6</sup> protoplasts per ml.

15-60 µg of plasmid DNA is mixed with 0.9 ml of protoplasts. The resulting suspension is mixed with 40% polyethylene glycol (MW 8000, PEG 8000), by gentle inversion a few times at room temperature for 5 to 25 min. Protoplast culture medium known in the art is added into the PEG-DNA-protoplast mixture. Protoplasts are incubated in the culture medium for 24 hour to 5 days and cell extracts can be used for assay of transient expression of the introduced gene. Alternatively, transformed cells can be used to produce transgenic callus, which in turn can be used to produce transgenic plants, by methods known in the art. See, for example, Nomura and Komamine, *Plt. Phys.* 79:988-991 (1985), *Identification and Isolation of Single Cells that Produce Somatic Embryos in Carrot Suspension Cultures.*

The invention being thus described, it will be apparent to one of ordinary skill in the art that various modifications of the materials and methods for practicing the invention can be made. Such modifications are to be

109

considered within the scope of the invention as defined by the following claims.

Each of the references from the patent and periodical literature cited herein is hereby expressly incorporated in its entirety by such citation.

5

110

**REF TABLE 1**

Maximum Length Sequence:

related to:

Clone IDs:

5 9581

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 1

- Ceres seq\_id 1007546

- Alternative transcription start site(s) located in SEQ

10 ID NO 1:

-96,-51,25,27,29,55,64

(B) Polypeptide Sequence

15 - Pat. Appln. SEQ ID NO 2

- Ceres seq\_id 1007547

- Location of start within SEQ ID NO 1: at 1 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 2: at 32 aa.

20 (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 1

- gi No. 3879939

25 - Description:

- % Identity: 45.9

- Alignment Length: 61

- Location of Alignment in SEQ ID NO 2: from 48 to

108

(B) Polypeptide Sequence

30 - Pat. Appln. SEQ ID NO 3

- Ceres seq\_id 1007548

- Location of start within SEQ ID NO 1: at 100 nt.

35 (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 2

- gi No. 3879939

40 - Description:

- % Identity: 45.9

- Alignment Length: 61

- Location of Alignment in SEQ ID NO 3: from 15 to

45 75

(B) Polypeptide Sequence

50 - Pat. Appln. SEQ ID NO 4

- Ceres seq\_id 1007549

- Location of start within SEQ ID NO 1: at 121 nt.

111

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 3

- gi No. 3879939

- Description:

- % Identity: 45.9

- Alignment Length: 61

- Location of Alignment in SEQ ID NO 4: from 8 to 68

10 - (Ba) Polypeptide Activities: Similar to yeast membrane  
protein activities

15 Maximum Length Sequence:

related to:

Clone IDs:

9568

402131

20 (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 5

- Ceres seq\_id 1007583

- Alternative transcription start site(s) located in SEQ  
ID NO 5:

2,3,4,7,10,11,12,17,42,43,300,505

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 6

- Ceres seq\_id 1007584

- Location of start within SEQ ID NO 5: at 55 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- KH domain

- Location within SEQ ID NO 6: from 47 to 95 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 4

- gi No. 133940

- Description:

- % Identity: 75.1

- Alignment Length: 250

- Location of Alignment in SEQ ID NO 6: from 1 to

249

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 7

- Ceres seq\_id 1007585

- Location of start within SEQ ID NO 5: at 184 nt.

50

112

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- KH domain

- Location within SEQ ID NO 7: from 4 to 52 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 5

- gi No. 133940

- Description:

- % Identity: 75.1

- Alignment Length: 250

- Location of Alignment in SEQ ID NO 7: from 1 to

206

15 Maximum Length Sequence:

related to:

Clone IDs:

91769

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 8

- Ceres seq\_id 1008148

- Alternative transcription start site(s) located in SEQ  
ID NO 8:

-19,2,3,4,5,6,7,9,10,11,12,14

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 9

- Ceres seq\_id 1008149

- Location of start within SEQ ID NO 8: at 3 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 6

- gi No. 4539292

- Description:

- % Identity: 89.9

- Alignment Length: 181

- Location of Alignment in SEQ ID NO 9: from 25 to

40 203

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 10

- Ceres seq\_id 1008150

- Location of start within SEQ ID NO 8: at 75 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 7

- gi No. 4539292

113

- Description:
- % Identity: 89.9
- Alignment Length: 181
- Location of Alignment in SEQ ID NO 10: from 1 to

5 179

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 11
- Ceres seq\_id 1008151
- Location of start within SEQ ID NO 8: at 210 nt.

10

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 8
- gi No. 4539292
- Description:
- % Identity: 89.9
- Alignment Length: 181
- Location of Alignment in SEQ ID NO 11: from 1 to

20 134

(Ba) Polypeptide Activities: Similar to 40S Ribosomal protein activities, and glycine rich RNA binding protein activities.

25

## Maximum Length Sequence:

## related to:

30 Clone IDs:

8898

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 12
- Ceres seq\_id 1008334
- Alternative transcription start site(s) located in SEQ ID NO 12:

35

-12,29,30

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 13
- Ceres seq\_id 1008335
- Location of start within SEQ ID NO 12: at 2 nt.

40

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 9
- gi No. 3602948
- Description:
- % Identity: 43.7
- Alignment Length: 71

50

114

- Location of Alignment in SEQ ID NO 13: from 178 to 248

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 14
- Ceres seq\_id 1008336
- Location of start within SEQ ID NO 12: at 59 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 14: at 22 aa.

10

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 10
- gi No. 3602948
- Description:
- % Identity: 43.7
- Alignment Length: 71
- Location of Alignment in SEQ ID NO 14: from 159 to 229

20

(Ba) Polypeptide Activities: Plant specific gene, Chloroplast specific gene.

## Maximum Length Sequence:

## related to:

25 Clone IDs:

8286

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 15
- Ceres seq\_id 1008701
- Alternative transcription start site(s) located in SEQ ID NO 15:

30

-6,2,3,5,6,7,18,24,25,28,31,33,35,37,42,50

35

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 16
- Ceres seq\_id 1008702
- Location of start within SEQ ID NO 15: at 1 nt.

40

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 11
- gi No. 1083282
- Description:
- % Identity: 41.4
- Alignment Length: 106
- Location of Alignment in SEQ ID NO 16: from 23 to

50 121



115

- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 17  
- Ceres seq\_id 1008703  
- Location of start within SEQ ID NO 15: at 67 nt.

5

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 12  
- gi No. 1083282  
- Description:  
- % Identity: 41.4  
- Alignment Length: 106  
- Location of Alignment in SEQ ID NO 17: from 1 to

15 99

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 18  
- Ceres seq\_id 1008704  
- Location of start within SEQ ID NO 15: at 2 nt.

20

(Ba) Polypeptide Activities: Similar to cytochrome C oxidase activities.

25

Maximum Length Sequence:

related to:

Clone IDs:

7792

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 19  
- Ceres seq\_id 1009033  
- Alternative transcription start site(s) located in SEQ

ID NO 19:

2,374

35

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 20  
- Ceres seq\_id 1009004  
- Location of start within SEQ ID NO 19: at 48 nt.

40

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 13  
- gi No. 3582320  
- Description:  
- % Identity: 32.6  
- Alignment Length: 44  
- Location of Alignment in SEQ ID NO 20: from 51 to

50 93

116

- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 21  
- Ceres seq\_id 1009005  
- Location of start within SEQ ID NO 19: at 57 nt.

5

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 14  
- gi No. 3582320  
- Description:  
- % Identity: 32.6  
- Alignment Length: 44  
- Location of Alignment in SEQ ID NO 21: from 48 to

15 90

(Ba) Polypeptide Activities: Similar to Ring-H2 Zinc Finger Protein activities.

20

Maximum Length Sequence:

related to:

Clone IDs:

7337

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 22  
- Ceres seq\_id 1009345  
- Alternative transcription start site(s) located in SEQ

ID NO 22:

2

30

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 23  
- Ceres seq\_id 1009346  
- Location of start within SEQ ID NO 22: at 50 nt.  
- Location of Signal Peptide Cleavage Site within SEQ

ID NO 23: at 22 aa.

35

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 15  
- gi No. 3176705  
- Description:  
- % Identity: 53.3  
- Alignment Length: 75  
- Location of Alignment in SEQ ID NO 23: from 1 to

75

(B) Polypeptide Sequence

50

117

- Pat. Appln. SEQ ID NO 24
- Ceres seq\_id 1009347
- Location of start within SEQ ID NO 22: at 62 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 24: at 18 aa.

5

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 16
- gi No. 3176705
- Description:
- % Identity: 53.3
- Alignment Length: 75
- Location of Alignment in SEQ ID NO 24: from 1 to 71

15

(Ba) Polypeptide Activities: Arabidopsis specific gene, dicot

specific gene.

specific gene, plant

20

Maximum Length Sequence:

related to:

Clone IDs:

6349

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 25
- Ceres seq\_id 1010140
- Alternative transcription start site(s) located in SEQ ID NO 25:  
-31, -29, 4, 5, 6, 10, 17, 34, 41, 749

30

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 26
- Ceres seq\_id 1010141
- Location of start within SEQ ID NO 25: at 3 nt.

35

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Bacterial regulatory proteins, deoR family
- Location within SEQ ID NO 26: from 57 to 95 aa.

40

(Dp) Related Amino Acid Sequences

- Alignment No. 17
- gi No. 3257798
- Description:
- % Identity: 40.6
- Alignment Length: 234
- Location of Alignment in SEQ ID NO 26: from 56 to 289

50

118

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 27
- Ceres seq\_id 1010142
- Location of start within SEQ ID NO 25: at 42 nt.

5

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Bacterial regulatory proteins, deoR family
- Location within SEQ ID NO 27: from 44 to 82 aa.

10

(Dp) Related Amino Acid Sequences

- Alignment No. 18
- gi No. 3257798
- Description:
- % Identity: 40.6
- Alignment Length: 234
- Location of Alignment in SEQ ID NO 27: from 43 to 276

20

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 28
- Ceres seq\_id 1010143
- Location of start within SEQ ID NO 25: at 231 nt.

25

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 19
- gi No. 3257798
- Description:
- % Identity: 40.6
- Alignment Length: 234
- Location of Alignment in SEQ ID NO 28: from 1 to 213

35

Maximum Length Sequence:  
related to:

Clone IDs:

6261

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 29
- Ceres seq\_id 1010217
- Alternative transcription start site(s) located in SEQ ID NO 29:  
2, 5, -5

45

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 30
- Ceres seq\_id 1010218
- Location of start within SEQ ID NO 29: at 35 nt.

50

119

- Location of Signal Peptide Cleavage Site within SEQ ID NO 30: at 22 aa.

5 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 20
- gi No. 3341723
- Description:
- % Identity: 64.3
- Alignment Length: 131
- Location of Alignment in SEQ ID NO 30: from 1 to

118

15 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 31

- Ceres seq id 1010219

- Location of start within SEQ ID NO 29: at 113 nt.

- Location of Signal Peptide Cleavage Site within SEQ ID NO 31: at 13 aa.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 21
- gi No. 3341723
- Description:
- % Identity: 64.3
- Alignment Length: 131
- Location of Alignment in SEQ ID NO 31: from 1 to

107

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 32

- Ceres seq id 1010220

- Location of start within SEQ ID NO 29: at 121 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 22
- gi No. 3341723
- Description:
- % Identity: 64.3
- Alignment Length: 131
- Location of Alignment in SEQ ID NO 32: from 1 to

106

(Ba) Polypeptide Activities: Similar to Constans like Protein activities and Zinc Finger Protein Activities.

120

Maximum Length Sequence:  
related to:

Clone IDs:  
6145

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 33
- Ceres seq id 1010302
- Alternative transcription start site(s) located in SEQ ID NO 33:  
-5,-3,-2,-1,2,3,4,5,6,7,8,10,11,12,13,15,19,23,45,349

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 34

- Ceres seq id 1010303

- Location of start within SEQ ID NO 33: at 59 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Pathogenesis-related protein Bet v 2 family
- Location within SEQ ID NO 34: from 5 to 155 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 23
- gi No. 1321731
- Description:
- % Identity: 35.7
- Alignment Length: 159
- Location of Alignment in SEQ ID NO 34: from 5 to

5 to

Maximum Length Sequence:  
related to:

Clone IDs:

5180

Public Genomic DNA:

gi No: 4757410

Predicted Exons:

INTR 37202 ... 37397

INTR 37493 ... 37825

OCXHAMG-CDNA

OCXHAMG-CDNA

INIT 37271 ... 37397

TERM 37493 ... 37704

OCXHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 35
- Ceres seq id 1010815
- Alternative transcription start site(s) located in SEQ ID NO 35:  
15,16,17,18,19,29,31,34

(B) Polypeptide Sequence

121

- Pat. Appln. SEQ ID NO 36
- Ceres seq\_id 1010816
- Location of start within SEQ ID NO 35: at 70 nt.

5

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- (Dp) Related Amino Acid Sequences
  - Alignment No. 24
  - gi No. 2879311
  - Description:
  - % Identity: 88.4
  - Alignment Length: 112
  - Location of Alignment in SEQ ID NO 36: from 1 to

112

15

(D) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 37
- Ceres seq\_id 1010817
- Location of start within SEQ ID NO 35: at 133 nt.

20

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- (Dp) Related Amino Acid Sequences
  - Alignment No. 25
  - gi No. 2879811
  - Description:
  - % Identity: 88.4
  - Alignment Length: 112
  - Location of Alignment in SEQ ID NO 37: from 1 to

30 91

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 38
- Ceres seq\_id 1010318
- Location of start within SEQ ID NO 35: at 257 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 38: at 26 aa.

35

(Ba) Polypeptide Activities: Similar to ribosomal protein activities.

40

Maximum Length Sequence:

related to:

Clone IDs:

42842

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 39
- Ceres seq\_id 1011437
- Alternative transcription start site(s) located in SEQ ID NO 39:

50

-25

122

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 40
  - Ceres seq\_id 1011438
  - Location of start within SEQ ID NO 39: at 2 nt.

5

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 41
  - Ceres seq\_id 1011439
  - Location of start within SEQ ID NO 39: at 1 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 41: at 29 aa.

10

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- (Dp) Related Amino Acid Sequences
  - Alignment No. 26
  - gi No. 3334271
  - Description:
  - % Identity: 29.6
  - Alignment Length: 54
  - Location of Alignment in SEQ ID NO 41: from 13 to

20

65

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 42
- Ceres seq\_id 1011440
- Location of start within SEQ ID NO 39: at 28 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 42: at 20 aa.

25

30

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- (Dp) Related Amino Acid Sequences
  - Alignment No. 27
  - gi No. 3334271
  - Description:
  - % Identity: 29.6
  - Alignment Length: 54
  - Location of Alignment in SEQ ID NO 42: from 4 to

35

56

(Ba) Polypeptide Activities: Similar to NADH Oxidoreductase  
MFE Subunit Protein Activities.

45

Maximum Length Sequence:

related to:

Clone IDs:

42475

(Ac) cDNA Polynucleotide Sequence

50

123

- Pat. Appln. SEQ ID NO 43
- Ceres seq\_id 1011616
- Alternative transcription start site(s) located in SEQ ID NO 43:

5 3,5,476

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 44
- Ceres seq\_id 1011617
- Location of start within SEQ ID NO 43: at 115 nt.

10

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 28
- gi No. 3859560
- Description:
- % Identity: 36.2
- Alignment Length: 225
- Location of Alignment in: SEQ ID NO 44: from 31 to 248

20

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 45
- Ceres seq\_id 1011618
- Location of start within SEQ ID NO 43: at 406 nt.

25

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 29
- gi No. 3859560
- Description:
- % Identity: 36.2
- Alignment Length: 225
- Location of Alignment in SEQ ID NO 45: from 1 to 151

35

- (Ba) Polypeptide Activities: Similar to acyl-protein thioesterases protein activities, calcium independent phospholipase A2 activities, and carboxylesterase activities.

40

45 Maximum Length Sequence:  
related to:

## Clone IDs:

42405

- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 46
- Ceres seq\_id 1011631

50

124

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 47
- Ceres seq\_id 1011632
- Location of start within SEQ ID NO 46: at 3 nt.

5

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 30
- gi No. 3618318
- Description:
- % Identity: 72.7
- Alignment Length: 44
- Location of Alignment in SEQ ID NO 47: from 91 to 134

15

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 48
- Ceres seq\_id 1011633
- Location of start within SEQ ID NO 46: at 9 nt.

20

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 31
- gi No. 3618318
- Description:
- % Identity: 72.7
- Alignment Length: 44
- Location of Alignment in SEQ ID NO 48: from 89 to 132

30

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 49
- Ceres seq\_id 1011634
- Location of start within SEQ ID NO 46: at 15 nt.

35

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 32
- gi No. 3618318
- Description:
- % Identity: 72.7
- Alignment Length: 44
- Location of Alignment in SEQ ID NO 49: from 87 to 130

45

50

125

(Ba) Polypeptide Activities: Similar to Constants protein activities, and zinc finger protein activities.

Maximum Length Sequence:

related to:

Clone IDs:

42240

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 50

- Ceres seq\_id 1011714

10

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 51

- Ceres seq\_id 1011715

- Location of start within SEQ ID NO 50: at 2 nt.

15

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Bacterial mutT protein

- Location within SEQ ID NO 51: from 26 to 67 aa.

20

(Dp) Related Amino Acid Sequences

- Alignment No. 33

- gi No. 2129134

- Description:

- % Identity: 40.8

- Alignment Length: 121

- Location of Alignment in SEQ ID NO 51: from 12 to

30 131

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 52

- Ceres seq\_id 1011716

- Location of start within SEQ ID NO 50: at 14 nt.

35

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Bacterial mutT protein

- Location within SEQ ID NO 52: from 22 to 63 aa.

40

(Dp) Related Amino Acid Sequences

- Alignment No. 34

- gi No. 2129134

- Description:

- % Identity: 40.8

- Alignment Length: 121

- Location of Alignment in SEQ ID NO 52: from 8 to

127

(B) Polypeptide Sequence

50

126

- Pat. Appln. SEQ ID NO 53

- Ceres seq\_id 1011717

- Location of start within SEQ ID NO 50: at 185 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 35

- gi No. 2129134

- Description:

- % Identity: 40.8

- Alignment Length: 121

- Location of Alignment in SEQ ID NO 53: from 1 to

70

Maximum Length Sequence:

related to:

Clone IDs:

42169

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 54

- Ceres seq\_id 1011784

- Alternative transcription start site(s) located in SEQ

ID NO 54:

-15, -4, 7, 402

25

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 55

- Ceres seq\_id 1011785

- Location of start within SEQ ID NO 54: at 1 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 55: at 42 aa.

30

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Plant lipid transfer protein family

- Location within SEQ ID NO 55: from 45 to 108 aa.

35

(Dp) Related Amino Acid Sequences

- Alignment No. 36

- gi No. 543565

- Description:

- % Identity: 56.5

- Alignment Length: 85

- Location of Alignment in SEQ ID NO 55: from 29 to

113

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 56

- Ceres seq\_id 1011786

- Location of start within SEQ ID NO 54: at 49 nt.

50

127

- Location of Signal Peptide Cleavage Site within SEQ ID NO 56: at 26 aa.

5

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- Plant lipid transfer protein family
  - Location within SEQ ID NO 56: from 29 to 92 aa.

10

- (Dp) Related Amino Acid Sequences
- Alignment No. 37
  - gi No. 543565
  - Description:
  - % Identity: 56.5
  - Alignment Length: 85
  - Location of Alignment in SEQ ID NO 56: from 13 to 94

15

(B)

Polypeptide Sequence

- Pat. Appln. SEQ ID NO 57
- Ceres seq\_id 1011787
- Location of start within SEQ ID NO 54: at 3 nt.

20

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

Maximum Length Sequence:  
related to:

Clone IDs:

41992

- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 58
  - Ceres seq\_id 1011820
  - Alternative transcription start site(s) located in SEQ ID NO 58:

35

-40,37

(B)

Polypeptide Sequence

- Pat. Appln. SEQ ID NO 59
- Ceres seq\_id 1011821
- Location of start within SEQ ID NO 58: at 1 nt.

40

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 38
- gi No. 3417418
- Description:
- % Identity: 23.6
- Alignment Length: 207

50

128

- Location of Alignment in SEQ ID NO 59: from 24 to 226

5

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 60
  - Ceres seq\_id 1011822
  - Location of start within SEQ ID NO 58: at 13 nt.

10

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 39
  - gi No. 3417418
  - Description:
  - % Identity: 23.6
  - Alignment Length: 207
  - Location of Alignment in SEQ ID NO 60: from 20 to 222

15

(B)

Polypeptide Sequence

- Pat. Appln. SEQ ID NO 61
- Ceres seq\_id 1011823
- Location of start within SEQ ID NO 58: at 151 nt.

20

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 40
- gi No. 3417418
- Description:
- % Identity: 23.6
- Alignment Length: 207
- Location of Alignment in SEQ ID NO 61: from 1 to 176

35

(Ba) Polypeptide Activities: Arabidopsis specific gene, dicot

specific gene.

specific gene, plant

Maximum Length Sequence:

related to:

Clone IDs:

41851

- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 62
  - Ceres seq\_id 1011874

50

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 63

129

- Ceres seq\_id 1011875
- Location of start within SEQ ID NO 62: at 1 nt.

5 Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences

- Alignment No. 41
- gi No. 2911044
- Description:
- % Identity: 78.7
- Alignment Length: 95

10 - Location of Alignment in SEQ ID NO 63: from 28 to 121

15 (Da) Polypeptide Activities: Arabidopsis specific gene,  
dicot specific gene, plant specific gene.

Maximum Length Sequence:

related to:  
Clone IDs:

20 41582

(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 64  
- Ceres seq\_id 1011981

(3) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 65
- Ceres seq\_id 1011982
- Location of start within SEQ ID NO 64: at 68 nt.

30 (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 42
- gi No. 4115355
- Description:
- % Identity: 100
- Alignment Length: 52
- Location of Alignment in SEQ ID NO 65: from 1 to

40 52

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 66
- Ceres seq\_id 1011983
- Location of start within SEQ ID NO 64: at 3 nt.

45 (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

(B) Polypeptide Sequence

50

130

- Pat. Appln. SEQ ID NO 67
- Ceres seq\_id 1011984
- Location of start within SEQ ID NO 64: at 483 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 67: at 19 aa.

10 (Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot specific gene, plant specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

15 33470

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 68
- Ceres seq\_id 1014547
- Alternative transcription start site(s) located in SEQ ID NO 68: -39, -2, -1, 2, 3, 7, 8, 9, 10

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 69
- Ceres seq\_id 1014548
- Location of start within SEQ ID NO 68: at 1 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 69: at 32 aa.

30 (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- Plant lipid transfer protein family
- Location within SEQ ID NO 69: from 44 to 110 aa.

35 (Dp) Related Amino Acid Sequences

- Alignment No. 43
- gi No. 306279:
- Description:
- % Identity: 72.2
- Alignment Length: 90
- Location of Alignment in SEQ ID NO 69: from 21 to

1:0

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 70
- Ceres seq\_id 1014549
- Location of start within SEQ ID NO 68: at 25 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 70: at 24 aa.

50



131

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Plant lipid transfer protein family
- Location within SEQ ID NO 70: from 36 to 102 aa.

5

## (Dp) Related Amino Acid Sequences

- Alignment No. 44
- gi No. 3062791
- Description: 72.2
- % Identity: 72.2
- Alignment Length: 90
- Location of Alignment in SEQ ID NO 70: from 13 to 102

10

## Maximum Length Sequence:

related to:

Clone IDs:

38004

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 71
- Ceres seq\_id 1014995
- Alternative transcription start site(s) located in SEQ ID NO 71:

2,3,4

25

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 72
- Ceres seq\_id 1014996
- Location of start within SEQ ID NO 71: at 2 nt.

30

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Aminotransferase class IV
- Location within SEQ ID NO 72: from 71 to 334 aa.

35

## (Dp) Related Amino Acid Sequences

- Alignment No. 45
- gi No. 3340183
- Description: 54.7
- % Identity: 54.7
- Alignment Length: 257
- Location of Alignment in SEQ ID NO 72: from 56 to 341

40

45

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 73
- Ceres seq\_id 1014997
- Location of start within SEQ ID NO 71: at 65 nt.

50

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

132

- Aminotransferase class IV
- Location within SEQ ID NO 73: from 50 to 313 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 46
- gi No. 3540183
- Description: 54.7
- % Identity: 54.7
- Alignment Length: 287
- Location of Alignment in SEQ ID NO 73: from 35 to 320

10

320

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 74
- Ceres seq\_id 1014998
- Location of start within SEQ ID NO 71: at 167 nt.

15

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Aminotransferase class IV
- Location within SEQ ID NO 74: from 16 to 279 aa.

20

## (Dp) Related Amino Acid Sequences

- Alignment No. 47
- gi No. 3540183
- Description: 54.7
- % Identity: 54.7
- Alignment Length: 287
- Location of Alignment in SEQ ID NO 74: from 1 to 286

25

30

## Maximum Length Sequence:

related to:

Clone IDs:

37701

Public Genomic DNA:

gi No: 4699904

Predicted Exons:

SINGLE 38530 ... 37988 OCKHBMG-CDS

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 75
- Ceres seq\_id 1015323
- Alternative transcription start site(s) located in SEQ ID NO 75:

-1,5,6,12

45

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 76
- Ceres seq\_id 1015324
- Location of start within SEQ ID NO 75: at 59 nt.

50

133

- Location of Signal Peptide Cleavage Site within SEQ ID NO 76: at 25 aa.

## (C) Nomination and Annotation of Domains within

## (Dp) Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 48
- gi No. 3860308
- Description:
- % Identity: 44.5
- Alignment Length: 140
- Location of Alignment in SEQ ID NO 76: from 56 to 75

## (Ba) Polypeptide Activities: Arabidopsis specific gene,

dicot

specific gene, plant

## Maximum Length Sequence:

related to:

Clone IDs:

364

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 77
- Ceres seq\_id 1016486
- Alternative transcription start site(s) located in SEQ ID NO 77: 17,19,20,21,22,23,29,35,38

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 78
- Ceres seq\_id 1016487
- Location of start within SEQ ID NO 77: at 1 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 78: at 46 aa.

## (C) Nomination and Annotation of Domains within

## (Dp) Predicted Polypeptide(s)

- Gamma-thionins family
- Location within SEQ ID NO 78: from 55 to 101 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 49
- gi No. 4038039
- Description:
- % Identity: 100
- Alignment Length: 77
- Location of Alignment in SEQ ID NO 78: from 25 to 101

134

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 79
- Ceres seq\_id 1016488
- Location of start within SEQ ID NO 77: at 73 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 79: at 22 aa.

## (C) Nomination and Annotation of Domains within

## (Dp) Predicted Polypeptide(s)

- Gamma-thionins family
- Location within SEQ ID NO 79: from 31 to 77 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 50
- gi No. 4038039
- Description:
- % Identity: 100
- Alignment Length: 77
- Location of Alignment in SEQ ID NO 79: from 1 to 77

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 80
- Ceres seq\_id 1016489
- Location of start within SEQ ID NO 77: at 118 nt.

## (C) Nomination and Annotation of Domains within

## (Dp) Predicted Polypeptide(s)

- Gamma-thionins family
- Location within SEQ ID NO 80: from 16 to 62 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 51
- gi No. 4038039
- Description:
- % Identity: 100
- Alignment Length: 77
- Location of Alignment in SEQ ID NO 80: from 1 to 62

## Maximum Length Sequence:

related to:

Clone IDs:

33891

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 81
- Ceres seq\_id 1018341
- Alternative transcription start site(s) located in SEQ ID NO 81: 4

135

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 82
- Ceres seq\_id 1018342
- Location of start within SEQ ID NO 81: at 71 nt.

5

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 52
- gi No. 1064887
- Description:
- % Identity: 60
- Alignment Length: 65
- Location of Alignment in SEQ ID NO 82: from 2 to

15 66

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 83
- Ceres seq\_id 1018343
- Location of start within SEQ ID NO 83: at 143 nt.

20

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 53
- gi No. 1064887
- Description:
- % Identity: 60
- Alignment Length: 65
- Location of Alignment in SEQ ID NO 83: from 1 to

30 42

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 84
- Ceres seq\_id 1018344
- Location of start within SEQ ID NO 81: at 146 nt.

35

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 54
- gi No. 1064887
- Description:
- % Identity: 60
- Alignment Length: 65
- Location of Alignment in SEQ ID NO 84: from 1 to

45

41

(Ba) Polypeptide Activities: Similar to pollen coat protein activities and LEA protein activities.

50

136

## Maximum Length Sequence:

related to:

Clone IDs:

33828

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 85
- Ceres seq\_id 1018382
- Alternative transcription start site(s) located in SEQ

ID NO 85:

10 2,4,5,6,7,9,10,11,14

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 86
- Ceres seq\_id 1018383
- Location of start within SEQ ID NO 85: at 22 nt.
- Location of Signal Peptide Cleavage Site within SEQ

ID NO 86: at 24 aa.

15

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Plant lipid transfer protein family
- Location within SEQ ID NO 86: from 28 to 115 aa.

20

## (Dp) Related Amino Acid Sequences

- Alignment No. 55
- gi No. 899224
- Description:
- % Identity: 78.2
- Alignment Length: 119
- Location of Alignment in SEQ ID NO 86: from 1 to

25

119

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 87
- Ceres seq\_id 1018384
- Location of start within SEQ ID NO 85: at 73 nt.

35

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Plant lipid transfer protein family
- Location within SEQ ID NO 87: from 11 to 98 aa.

40

## (Dp) Related Amino Acid Sequences

- Alignment No. 56
- gi No. 899224
- Description:
- % Identity: 78.2
- Alignment Length: 119
- Location of Alignment in SEQ ID NO 87: from 1 to

45

50 102

137

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 88
- Ceres seq\_id 1018385
- Location of start within SEQ ID NO 85: at 3 nt.

5

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

10 Maximum Length Sequence:

related to:

Clone IDs:

30349

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 89
- Ceres seq\_id 1020666
- Alternative transcription start site(s) located in SEQ ID NO 89:  
33, 35, 39, 40, 42, 43, 44, 45, 64, 173

20

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 90
- Ceres seq\_id 1020667
- Location of start within: SEQ ID NO 89: at 118 nt.

25

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein S7e
- Location within SEQ ID NO 90: from 7 to 187 aa.

30

## (Dp) Related Amino Acid Sequences

- Alignment No. 57

- gi No. 3851636

- Description:

- % Identity: 77.4

- Alignment Length: 190

- Location of Alignment in SEQ ID NO 90: from 1 to

190

40

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 91
- Ceres seq\_id 1020668
- Location of start within SEQ ID NO 89: at 271 nt.

45

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein S7e
- Location within SEQ ID NO 91: from 1 to 136 aa.

50

## (Dp) Related Amino Acid Sequences

- Alignment No. 58

138

- gi No. 3851636
- Description:
- % Identity: 77.4
- Alignment Length: 190
- Location of Alignment in SEQ ID NO 91: from 1 to

5

139

Maximum Length Sequence:

related to:

Clone IDs:

30113

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 92
- Ceres seq\_id 1020784

15

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 93
- Ceres seq\_id 1020785
- Location of start within SEQ ID NO 92: at 60 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 93: at 25 aa.

20

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 59
- gi No. 3860308
- Description:
- % Identity: 44.5
- Alignment Length: 140
- Location of Alignment in SEQ ID NO 93: from 56 to

30

175

(Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot specific gene, plant specific gene.

35

40 Maximum Length Sequence:

related to:

Clone IDs:

29123

Public Genomic DNA:

gi No: 5822667

Predicted Exons:

INTR 68772 ... 69532 OCKHAMG-CDNA

SINGLE 68846 ... 69325 OCKHAMG-CDS

gi No: 6041831

Predicted Exons:

- INTR 63702 ... 64462 OCKHAMG-CDNA 139
- SINGLE 63776 ... 64255 OCKHAMG-CDS
- (Ac) cDNA Polynucleotide Sequence
- 5 - Pat. Appln. SEQ ID NO 94
- Ceres seq\_id 1021525
- Alternative transcription start site(s) located in SEQ ID NO 94: 25,26,27,28,29,35,36,39,51,53,54,68
- 10 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 95
- Ceres seq\_id 1021526
- Location of start within SEQ ID NO 94: at 75 nt.
- 15 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 60
- gi No. 4388983
- Description:
- % Identity: 29.1
- Alignment Length: 110
- Location of Alignment in SEQ ID NO 95: from 47 to 156
- 25 (Ba) Polypeptide Activities: Similar to adrenodoxin precursor protein activities and adrenal ferredoxin activities.
- 30 Maximum Length Sequence: related to:
- Clone IDs: 2891
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 96
- Ceres seq\_id 1021563
- Alternative transcription start site(s) located in SEQ ID NO 96: 16,28,29,30,31,35,36,43,74,77,80,88,89,90,95
- 40 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 97
- Ceres seq\_id 1021564
- Location of start within SEQ ID NO 96: at 2 nt.
- 45 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 61
- 50

- 140
- gi No. 2493089
- Description:
- % Identity: 29.7
- Alignment Length: 77
- Location of Alignment in SEQ ID NO 97: from 52 to 125
- 5 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 98
- Ceres seq\_id 1021565
- Location of start within SEQ ID NO 96: at 116 nt.
- 10 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 62
- gi No. 2493089
- Description:
- % Identity: 29.7
- Alignment Length: 77
- Location of Alignment in SEQ ID NO 98: from 14 to 87
- 20 (Ba) Polypeptide Activities: Similar to ATPK-mouse activities, ATP synthase activities, and mitochondrial F-Chain activities.
- 25 Maximum Length Sequence: related to:
- Clone IDs: 28979
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 99
- Ceres seq\_id 1021576
- Alternative transcription start site(s) located in SEQ ID NO 99: -13,-
- 35 11,3,5,7,11,12,13,15,24,26,28,29,31,34,35,36,37,38,39
- 40 43,44,45,46,47,48,50,51,52,56,58,59,61,67,68,70,75,82,83,91,112
- 244,313,318
- 45 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 100
- Ceres seq\_id 1021577
- Location of start within SEQ ID NO 99: at 109 nt.
- 50 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- Photosystem I psaG / psaK

141

- Location within SEQ ID NO 100: from 50 to 130 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 63
- gi No. 3885511
- Description:
- % Identity: 81.1
- Alignment Length: 128
- Location of Alignment in SEQ ID NO 100: from 1 to

127

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 101
- Ceres seq\_id 1021578
- Location of start within SEQ ID NO 99: at 121 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Photosystem I psaG / psaK
- Location within SEQ ID NO 101: from 46 to 126 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 64
- gi No. 3885511
- Description:
- % Identity: 81.1
- Alignment Length: 128
- Location of Alignment in SEQ ID NO 101: from 1 to

123

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 102
- Ceres seq\_id 1021579
- Location of start within SEQ ID NO 99: at 124 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Photosystem I psaG / psaK
- Location within SEQ ID NO 102: from 45 to 125 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 65
- gi No. 3885511
- Description:
- % Identity: 81.1
- Alignment Length: 128
- Location of Alignment in SEQ ID NO 102: from 1 to

122

- 0 Maximum Length Sequence:  
related to:

142

## Clone IDs:

28177

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 103
- Ceres seq\_id 1021927
- Alternative transcription start site(s) located in SEQ ID NO 103:  
2, 12, 20, 23, 29, 36, 46, 47, 53

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 104
- Ceres seq\_id 1021928
- Location of start within SEQ ID NO 103: at 67 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 66
- gi No. 4263779
- Description:
- % Identity: 28.2
- Alignment Length: 177
- Location of Alignment in SEQ ID NO 104: from 14 to 182

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 105
- Ceres seq\_id 1021929
- Location of start within SEQ ID NO 103: at 172 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 105: at 23 aa.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 67
- gi No. 4263779
- Description:
- % Identity: 28.2
- Alignment Length: 177
- Location of Alignment in SEQ ID NO 105: from 1 to

147

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 106
- Ceres seq\_id 1021930
- Location of start within SEQ ID NO 103: at 220 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

143

- 5                   - Alignment No. 68  
                  - gi No. 4263779  
                  - Description:  
                  - % Identity: 28.2  
                  - Alignment Length: 177  
                  - Location of Alignment in SEQ ID NO 106: from 1 to 131

10           (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot specific gene, plant specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

2807

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 107
- Ceres seq\_id 1021945

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 108
- Ceres seq\_id 1021946
- Location of start within SEQ ID NO 107: at 1 nt.

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 69
- gi No. 3334132
- Description:
- % Identity: 27.4
- Alignment Length: 114
- Location of Alignment in SEQ ID NO 108: from 28 to 136

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 109
- Ceres seq\_id 1021947
- Location of start within SEQ ID NO 107: at 163 nt.

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 70
- gi No. 3334132
- Description:
- % Identity: 27.4
- Alignment Length: 114
- Location of Alignment in SEQ ID NO 109: from 1 to 82

144

(Ba) Polypeptide Activities: Similar to C21 ORF4 - membrane protein activities

5           Maximum Length Sequence:

related to:

Clone IDs:

27792

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 110
- Ceres seq\_id 1022170
- Alternative transcription start site(s) located in SEQ ID NO 110: -4, -1, 32, 68

15           (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 111
- Ceres seq\_id 1022171
- Location of start within SEQ ID NO 110: at 92 nt.

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 71
- gi No. 1173456
- Description:
- % Identity: 54.7
- Alignment Length: 129
- Location of Alignment in SEQ ID NO 111: from 4 to 131

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 112
- Ceres seq\_id 1022172
- Location of start within SEQ ID NO 110: at 191 nt.

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 72
- gi No. 1173456
- Description:
- % Identity: 54.7
- Alignment Length: 129
- Location of Alignment in SEQ ID NO 112: from 1 to 98

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 113
- Ceres seq\_id 1022173
- Location of start within SEQ ID NO 110: at 1 nt.

145

- Location of Signal Peptide Cleavage Site within SEQ ID NO 113: at 19 aa.

5 (Ba) Polypeptide Activities: Similar to small nuclear ribonucleoprotein activities.

Maximum Length Sequence:

related to:

Clone IDs:

27167

Public Genomic DNA:

gi No: 3046830

Predicted Exons:

INIT 45217 ... 45131 OCKHANG-CDS

INTR 44595 ... 44629 OCKHANG-CDS

TERM 44554 ... 44286 OCKHANG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 114

- Ceres seq\_id 1022554

20 - Alternative transcription start site(s) located in SEQ ID NO 114:

-49,-

85

19,7,22,23,28,29,30,31,32,33,35,36,37,51,57,67,75,81,82

25

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 115

- Ceres seq\_id 1022555

- Location of start within SEQ ID NO 114: at 107 nt.

30

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Heme-binding domain in cytochrome b5 and

oxidoreductases

35 - Location within SEQ ID NO 115: from 7 to 84 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 73

- gi No. 4240122

- Description:

- % Identity: 100

- Alignment Length: 140

- Location of Alignment in SEQ ID NO 115: from 1 to

140

45

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 116

- Ceres seq\_id 1022556

- Location of start within SEQ ID NO 114: at 317 nt.

50

146

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 74

- gi No. 4240122

- Description:

- % Identity: 100

- Alignment Length: 140

- Location of Alignment in SEQ ID NO 116: from 1 to

10 70

Maximum Length Sequence:

related to:

Clone IDs:

27109

Public Genomic DNA:

gi No: 6449507

Predicted Exons:

INIT 94711 ... 94519 OCKHANG-CDS

INTR 94417 ... 94326 OCKHANG-CDS

INTR 94249 ... 94131 OCKHANG-CDS

TERM 94046 ... 93968 OCKHANG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 117

- Ceres seq\_id 1022594

25 - Alternative transcription start site(s) located in SEQ ID NO 117:

2,15,24,25,66,69,72,74

30

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 118

- Ceres seq\_id 1022595

- Location of start within SEQ ID NO 117: at 1 nt.

35

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Universal stress protein family

- Location within SEQ ID NO 118: from 98 to 163 aa.

40

(Dp) Related Amino Acid Sequences

- Alignment No. 75

- gi No. 2160182

- Description:

- % Identity: 37.5

- Alignment Length: 160

- Location of Alignment in SEQ ID NO 118: from 39

to 186

45

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 119

- Ceres seq\_id 1022596

50



147

- 5 - Location of start within SEQ ID NO 117: at 82 nt.
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- Universal stress protein family
  - Location within SEQ ID NO 119: from 71 to 156 aa.

## (Dp) Related Amino Acid Sequences

- 10 - Alignment No. 76
- gi No. 2160182
  - Description:
  - % Identity: 37.5
  - Alignment Length: 160
  - Location of Alignment in SEQ ID NO 119: from 12 to 159

## (B) Polypeptide Sequence

- 20 - Pat. Appln. SEQ ID NO 120
- Ceres seq\_id 1022597
  - Location of start within SEQ ID NO 117: at 106 nt.

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- 25 - Universal stress protein family
- Location within SEQ ID NO 120: from 63 to 148 aa.

## (Dp) Related Amino Acid Sequences

- 30 - Alignment No. 77
- gi No. 2160182
  - Description:
  - % Identity: 37.5
  - Alignment Length: 160
  - Location of Alignment in SEQ ID NO 120: from 4 to 151

- 35 (Ba) Polypeptide Activities: Similar to protein in methanobacterium thermoautotrophicum activities.

## Maximum Length Sequence:

related to:

Clone IDs:

26994

## (Ac) cDNA Polynucleotide Sequence

- 45 - Pat. Appln. SEQ ID NO 121
- Ceres seq\_id 1022621
  - Alternative transcription start site(s) located in SEQ ID NO 121: 2,7,9,13,35,38,45,57

- 50 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 122

148

- Ceres seq\_id 1022622
- Location of start within SEQ ID NO 121: at 86 nt.

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- 5 (Dp) Related Amino Acid Sequences
- Alignment No. 78
  - gi No. 3256599
  - Description:
  - % Identity: 32
  - Alignment Length: 128
  - Location of Alignment in SEQ ID NO 122: from 7 to 128

- 10 - Alignment No. 78

- 15 (Ba) Polypeptide Activities: Similar to structural cell wall protein activities, and larval gene protein in the Fruit fly activities.

## Maximum Length Sequence:

related to:

Clone IDs:

23518

## (Ac) cDNA Polynucleotide Sequence

- 25 - Pat. Appln. SEQ ID NO 123
- Ceres seq\_id 1024375
  - Alternative transcription start site(s) located in SEQ ID NO 123: 7,8,10,17,18,24,35,41,42,43,44,46,48,52,54,59

## (B) Polypeptide Sequence

- 30 - Pat. Appln. SEQ ID NO 124
- Ceres seq\_id 1024376
  - Location of start within SEQ ID NO 123: at 130 nt.

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- 35 - Ribosomal protein S7e
- Location within SEQ ID NO 124: from 7 to 137 aa.

## (Dp) Related Amino Acid Sequences

- 40 - Alignment No. 79
- gi No. 3851636
  - Description:
  - % Identity: 76.6
  - Alignment Length: 188
  - Location of Alignment in SEQ ID NO 124: from 1 to 188

- 50 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 125

149

- 5 Ceres seq\_id 1024377  
- Location of start within SEQ ID NO 123: at 283 nt.
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
- Ribosomal protein S7e  
- Location within SEQ ID NO 125: from 1 to 136 aa.

## (Dp) Related Amino Acid Sequences

- 10 - Alignment No. 80  
- gi No. 3851636  
- Description:  
- % Identity: 76.6  
- Alignment Length: 188  
- Location of Alignment in SEQ ID NO 125: from 1 to 136 aa.

137

## Maximum Length Sequence:

related to:

Clone IDs:

23170

## (Ac) cDNA Polynucleotide Sequence

- 20 - Pat. Appln. SEQ ID NO 126  
- Ceres seq\_id 1024535  
- Alternative transcription start site(s) located in SEQ ID NO 126:  
2,8,11,31,46,47,48

## (B) Polypeptide Sequence

- 30 - Pat. Appln. SEQ ID NO 127  
- Ceres seq\_id 1024536  
- Location of start within SEQ ID NO 126: at 115 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 35 - Alignment No. 81  
- gi No. 2621731  
- Description:  
- % Identity: 35.2  
- Alignment Length: 88  
- Location of Alignment in SEQ ID NO 127: from 5 to 92

## (B) Polypeptide Sequence

- 45 - Pat. Appln. SEQ ID NO 128  
- Ceres seq\_id 1024537  
- Location of start within SEQ ID NO 126: at 3 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

150

## (Dp) Related Amino Acid Sequences

- 5 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 129  
- Ceres seq\_id 1024538  
- Location of start within SEQ ID NO 126: at 253 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 10 - Alignment No. 82  
- gi No. 2621731  
- Description:  
- % Identity: 35.2  
- Alignment Length: 88  
- Location of Alignment in SEQ ID NO 129: from 1 to 46

## (Ba) Polypeptide Activities: Similar to small nuclear

ribonucleoprotein activities.

## Maximum Length Sequence:

related to:

Clone IDs:

21228

## Public Genomic DNA:

gi No: 4139402

## Predicted Exons:

- 30 INIT 36726 ... 36352 GENBANK  
INTR 36257 ... 36135 GENBANK  
TERM 35555 ... 35469 GENBANK

- 35 INTR 36791 ... 36352 OCKHANG-CDNA  
INTR 36257 ... 36135 OCKHANG-CDNA  
INTR 35555 ... 35325 OCKHANG-CDNA

- 40 INIT 36726 ... 36352 OCKHANG-CDS  
INTR 36257 ... 36135 OCKHANG-CDS  
TERM 35555 ... 35469 OCKHANG-CDS

## (Ac) cDNA Polynucleotide Sequence

- 45 - Pat. Appln. SEQ ID NO 130  
- Ceres seq\_id 1025683  
- Alternative transcription start site(s) located in SEQ ID NO 130:  
2,8,22

## (B) Polypeptide Sequence

- 50 - Pat. Appln. SEQ ID NO 131  
- Ceres seq\_id 1025684  
- Location of start within SEQ ID NO 130: at 3 nt.

151

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- 5 - Ribosomal protein L6  
- Location within SEQ ID NO 131: from 34 to 216 aa.

(Dp) Related Amino Acid Sequences

- 10 - Alignment No. 83  
- gi No. 266945  
- Description:  
- % Identity: 84  
- Alignment Length: 194  
- Location of Alignment in SEQ ID NO 131: from 23 to 216

(B) Polypeptide Sequence

- 15 - Pat. Appln. SEQ ID NO 132  
- Ceres seq\_id 1025685  
- Location of start within SEQ ID NO 130: at 69 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- 20 - Ribosomal protein L6  
- Location within SEQ ID NO 132: from 12 to 194 aa.

(Dp) Related Amino Acid Sequences

- 25 - Alignment No. 84  
- gi No. 266945  
- Description:  
- % Identity: 84  
- Alignment Length: 194  
- Location of Alignment in SEQ ID NO 132: from 1 to 194

(B) Polypeptide Sequence

- 35 - Pat. Appln. SEQ ID NO 133  
- Ceres seq\_id 1025686  
- Location of start within SEQ ID NO 130: at 96 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- 40 - Ribosomal protein L6  
- Location within SEQ ID NO 133: from 3 to 185 aa.

(Dp) Related Amino Acid Sequences

- 45 - Alignment No. 85  
- gi No. 266945  
- Description:  
- % Identity: 84  
- Alignment Length: 194

152

- 185 - Location of Alignment in SEQ ID NO 133: from 1 to 185

Maximum Length Sequence:

related to:

Clone IDs:

19274

(Ac) cDNA Polynucleotide Sequence

- 10 - Pat. Appln. SEQ ID NO 134  
- Ceres seq\_id 1027152  
- Alternative transcription start site(s) located in SEQ ID NO 134:  
-350,3,4,10,11,-3,222

(B) Polypeptide Sequence

- 15 - Pat. Appln. SEQ ID NO 135  
- Ceres seq\_id 1027153  
- Location of start within SEQ ID NO 134: at 3 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- 20 - Alignment No. 86  
- gi No. 2879811  
- Description:  
- % Identity: 86.6  
- Alignment Length: 112  
- Location of Alignment in SEQ ID NO 135: from 26 to 137

(B) Polypeptide Sequence

- 30 - Pat. Appln. SEQ ID NO 136  
- Ceres seq\_id 1027154  
- Location of start within SEQ ID NO 134: at 78 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- 35 - Alignment No. 87  
- gi No. 2879811  
- Description:  
- % Identity: 86.6  
- Alignment Length: 112  
- Location of Alignment in SEQ ID NO 136: from 1 to 112

(B) Polypeptide Sequence

- 45 - Pat. Appln. SEQ ID NO 137  
- Ceres seq\_id 1027155  
- Location of start within SEQ ID NO 134: at 141 nt.

- 153  
(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 88  
- gi No. 2879811  
- Description:  
- % Identity: 86.6  
- Alignment Length: 112  
- Location of Alignment in SEQ ID NO 137: from 1 to 10 91
- (Ba) Polypeptide Activities: Similar to ribosomal L30 protein activities.
- 15 Maximum Length Sequence:  
related to:  
Clone IDs:  
17835
- 20 (Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 138  
- Ceres seq\_id 1028095  
- Alternative transcription start site(s) located in SEQ ID NO 138:  
-2,2,3,4,5,6,12,14,18,22,26,40,42,44,45,46,47
- 25 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 139  
- Ceres seq\_id 1028096  
- Location of start within SEQ ID NO 138: at 55 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 139: at 29 aa.
- 30 (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 89  
- gi No. 4336325  
- Description:  
- % Identity: 31.7  
- Alignment Length: 126  
- Location of Alignment in SEQ ID NO 139: from 15 to 135
- 45 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 140  
- Ceres seq\_id 1028097  
- Location of start within SEQ ID NO 138: at 214 nt.
- 50 (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- 154  
(Dp) Related Amino Acid Sequences  
- Alignment No. 90  
- gi No. 4336325  
- Description:  
- % Identity: 31.7  
- Alignment Length: 126  
- Location of Alignment in SEQ ID NO 140: from 1 to 82
- 10 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 141  
- Ceres seq\_id 1028098  
- Location of start within SEQ ID NO 138: at 3 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 141: at 16 aa.
- 15 (Ba) Polypeptide Activities: Similar to human C214 membrane protein activities.
- 20 Maximum Length Sequence:  
related to:  
Clone IDs:  
17075
- 25 (Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 142  
- Ceres seq\_id 1028608  
- Alternative transcription start site(s) located in SEQ ID NO 142:  
-4,2,28,31,36,49,59
- 30 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 143  
- Ceres seq\_id 1028609  
- Location of start within SEQ ID NO 142: at 95 nt.
- 35 (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 91  
- gi No. 2735528  
- Description:  
- % Identity: 33.9  
- Alignment Length: 118  
- Location of Alignment in SEQ ID NO 143: from 64 to 178
- 45 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 144  
- Ceres seq\_id 1028610  
- Location of start within SEQ ID NO 142: at 176 nt.

155

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 5 - Alignment No. 92  
- gi No. 2735528  
- Description:  
- % Identity: 33.9  
- Alignment Length: 118  
- Location of Alignment in SEQ ID NO 144: from 37 to 151

## (B) Polypeptide Sequence

- 15 - Pat. Appln. SEQ ID NO 145  
- Ceres seq\_id 102861  
- Location of start within SEQ ID NO 142: at 351 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 145: at 4 aa.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

## Maximum Length Sequence:

## related to:

## Clone IDs:

1505

## (Ac) cDNA Polynucleotide Sequence

- 30 - Pat. Appln. SEQ ID NO 146  
- Ceres seq\_id 1030069

## (B) Polypeptide Sequence

- 35 - Pat. Appln. SEQ ID NO 147  
- Ceres seq\_id 1030070  
- Location of start within SEQ ID NO 146: at 2 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 40 - Alignment No. 93  
- gi No. 4335755  
- Description:  
- % Identity: 63.6  
- Alignment Length: 143  
- Location of Alignment in SEQ ID NO 147: from 34 to 171

## (B) Polypeptide Sequence

- 50 - Pat. Appln. SEQ ID NO 148  
- Ceres seq\_id 1030071

156

- Location of start within SEQ ID NO 146: at 50 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 5 - Alignment No. 94  
- gi No. 4335755  
- Description:  
- % Identity: 63.6  
- Alignment Length: 143  
- Location of Alignment in SEQ ID NO 148: from 18 to 155

## (B) Polypeptide Sequence

- 15 - Pat. Appln. SEQ ID NO 149  
- Ceres seq\_id 1030072  
- Location of start within SEQ ID NO 145: at 170 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 20 - Alignment No. 95  
- gi No. 4335755  
- Description:  
- % Identity: 63.6  
- Alignment Length: 143  
- Location of Alignment in SEQ ID NO 149: from 1 to 115

- 30 (Ra) Polypeptide Activities: Similar to hydroxyproline-rich protein activities.

## Maximum Length Sequence:

## related to:

## Clone IDs:

12487

## (Ac) cDNA Polynucleotide Sequence

- 40 - Pat. Appln. SEQ ID NO 150  
- Ceres seq\_id 1032069  
- Alternative transcription start site(s) located in SEQ ID NO 150:  
4, 17

## (B) Polypeptide Sequence

- 45 - Pat. Appln. SEQ ID NO 151  
- Ceres seq\_id 1032070  
- Location of start within SEQ ID NO 150: at 74 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

157

## (Dp) Related Amino Acid Sequences

- Alignment No. 96
- gi No. 3386621
- Description:
- % Identity: 97.4
- Alignment Length: 234
- Location of Alignment in SEQ ID NO 151: from 1 to 231

5

## 10 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 152
- Ceres seq\_id 1032071
- Location of start within SEQ ID NO 150: at 122 nt.

## 15 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 97
- gi No. 3386621
- Description:
- % Identity: 97.4
- Alignment Length: 234
- Location of Alignment in SEQ ID NO 152: from 1 to 215

20

## 25 (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot specific gene, plant specific gene.

## Maximum length Sequence:

related to:

Clone IDs:

11466

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 153
- Ceres seq\_id 1033557
- Alternative transcription start site(s) located in SEQ ID NO 153: 62, 64, 65, 67, 72, 73, 74, 75, 166

35

## 40 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 154
- Ceres seq\_id 1033558
- Location of start within SEQ ID NO 153: at 94 nt.

## 45 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Ribosomal protein L34
- Location within SEQ ID NO 154: from 105 to 145

aa.

## (Dp) Related Amino Acid Sequences

50

158

- Alignment No. 98
- gi No. 132909
- Description:
- % Identity: 70
- Alignment Length: 30
- Location of Alignment in SEQ ID NO 154: from 116 to 145

5

## 10 (Ba) Polypeptide Activities: Similar to 50S ribosomal protein L34 activities.

## Maximum length Sequence:

related to:

Clone IDs:

21589

106951

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 155
- Ceres seq\_id 1034688
- Alternative transcription start site(s) located in SEQ ID NO 155: 2
- Clone 21589 starts at 2 and ends at in cDNA.

20

## 25 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 156
- Ceres seq\_id 1034689
- Location of start within SEQ ID NO 155: at 2 nt.

## 30 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 99
- gi No. 4335755
- Description:
- % Identity: 65
- Alignment Length: 143
- Location of Alignment in SEQ ID NO 156: from 33 to 170

35

## 40 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 157
- Ceres seq\_id 1034690
- Location of start within SEQ ID NO 155: at 47 nt.

## 45 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- (Dp) Related Amino Acid Sequences
- Alignment No. 100
- gi No. 4335755
- Description:

50

- 159
- % Identity: 65
  - Alignment Length: 143
  - Location of Alignment in SEQ ID NO 157: from 18 to 155

5

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 158
  - Ceres seq\_id 1034691
  - Location of start within SEQ ID NO 155: at 167 nt.

10

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 101
  - gi No. 4335755
  - Description:
  - % Identity: 65
  - Alignment Length: 143
  - Location of Alignment in SEQ ID NO 158: from 1 to 115

20

- (Ba) Polypeptide Activities: Similar to hydroxy proline rich glycoprotein activities.

25

Maximum Length Sequence:

related to:

Clone IDs:

10433

- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 159
  - Ceres seq\_id 1035033
  - Alternative transcription start site(s) located in SEQ ID NO 159:

ID NO 159:

32

35

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 160
  - Ceres seq\_id 1035034
  - Location of start within SEQ ID NO 159: at 3 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 160: at 47 aa.

40

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 102

- gi No. 3062795

- Description:

- % Identity: 39.8

- Alignment Length: 83

50

- 160
- Location of Alignment in SEQ ID NO 160: from 24 to 99

5

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 161
  - Ceres seq\_id 1035035
  - Location of start within SEQ ID NO 159: at 69 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 161: at 25 aa.

10

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 103
  - gi No. 3062795
  - Description:
  - % Identity: 39.8
  - Alignment Length: 83
  - Location of Alignment in SEQ ID NO 161: from 2 to 77

20

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 162
  - Ceres seq\_id 1035036
  - Location of start within SEQ ID NO 159: at 72 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 162: at 24 aa.

25

ID NO 162:

(C) Nomination and Annotation of Domains within

(Dp) Related Amino Acid Sequences

- Alignment No. 104

- gi No. 3062795

- Description:

- % Identity: 39.8

- Alignment Length: 83

- Location of Alignment in SEQ ID NO 162: from 1 to 76

35

- (Ba) Polypeptide Activities: Similar to Pollen coat protein activities.

40

Maximum Length Sequence:

related to:

Clone IDs:

10394

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 163

- Ceres seq\_id 1035071

50

161

- Alternative transcription start site(s) located in SEQ ID NO 163:
- 2,-1,2,3,18,19,21,24,25,56,66

- 5 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 164
  - Ceres seq\_id 1035072
  - Location of start within SEQ ID NO 163: at 66 nt.

- 10 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- KH domain
  - Location within SEQ ID NO 164: from 47 to 95 aa.

- 15 (Dp) Related Amino Acid Sequences

- Alignment No. 105
- gi No. 133940
- Description:
- % Identity: 76
- Alignment Length: 246
- Location of Alignment in SEQ ID NO 164: from 1 to 148

- 25 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 165
  - Ceres seq\_id 1035073
  - Location of start within SEQ ID NO 163: at 195 nt.

- 30 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- KH domain
  - Location within SEQ ID NO 165: from 4 to 52 aa.

- 35 (Dp) Related Amino Acid Sequences

- Alignment No. 106
- gi No. 133940
- Description:
- % Identity: 76
- Alignment Length: 246
- Location of Alignment in SEQ ID NO 165: from 1 to 105

- 45 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 166
  - Ceres seq\_id 1035074
  - Location of start within SEQ ID NO 163: at 513 nt.

- 50 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- (Dp) Related Amino Acid Sequences
- Alignment No. 107

162

- gi No. 133940
- Description:
- % Identity: 76
- Alignment Length: 246
- Location of Alignment in SEQ ID NO 166: from 1 to 97

- Maximum Length Sequence:  
related to:

- 10 Clone IDs:

10511

- Public Genomic DNA:

gi No: 4539290

- Predicted Exons:

INTR 5588 ... 5314 OCKHAMG-CDNA

INTR 5517 ... 5314 OCKHAMG-CDNA

gi No: 4914454

- Predicted Exons:

INTR 38937 ... 38663 OCKHAMG-CDNA

INTR 38866 ... 38663 OCKHAMG-CDNA

- (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 167

- Ceres seq\_id 1376587

- Alternative transcription start site(s) located in SEQ ID NO 167:
- 8,9,10,12,13,14,15,16,17,30,34,39,41

- 30 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 168

- Ceres seq\_id 1376588

- Location of start within SEQ ID NO 167: at 2 nt.

- 35 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 108

- gi No. 4539292

- Description:

- % Identity: 99.4

- Alignment Length: 177

- Location of Alignment in SEQ ID NO 168: from 26 to 202

- 45 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 169

- Ceres seq\_id 1376589

- Location of start within SEQ ID NO 167: at 77 nt.



163

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 5 - Alignment No. 109
- gi No. 4339292
- Description:
- % Identity: 99.4
- Alignment Length: 177
- Location of Alignment in SEQ ID NO 169: from 1 to

10 177

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 170
- Ceres seq\_id 1376590
- Location of start within SEQ ID NO 167: at 212 nt.

15

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 20 - Alignment No. 110
- gi No. 4539292
- Description:
- % Identity: 99.4
- Alignment Length: 177
- Location of Alignment in SEQ ID NO 170: From 1 to

25 132

## (3a) Polypeptide Activities: Similar 40S ribosomal protein activities.

## Maximum Length Sequence:

## related to:

## Clone IDs:

112-110

## Public Genomic DNA:

gi No: 4263774

## Predicted Exons:

- 40 INIT 1816 ... 1814 OCKHAMG-CDS
- INTR 1290 ... 1112 OCKHAMG-CDS
- TERM 959 ... 803 OCKHAMG-CDS

gi No: 4510360

## Predicted Exons:

- 45 INIT 114707... 114705 OCKHAMG-CDS
- INTR 114181... 114003 OCKHAMG-CDS
- TERM 113850... 113694 OCKHAMG-CDS

## (Ac) cDNA Polynucleotide Sequence:

- Pat. Appln. SEQ ID NO 171
- Ceres seq\_id 137858:
- Alternative transcription start site(s) located in SEQ

50 ID NO 171:

-35,-4,-3,18,20

164

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 172
- Ceres seq\_id 1378582
- Location of start within SEQ ID NO 171: at 112 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 172: at 17 aa.

5

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- DnaJ domain
- Location within SEQ ID NO 172: from 57 to 101 aa.

10

## (Dp) Related Amino Acid Sequences

- 15 - Alignment No. 111
- gi No. 4263775
- Description:
- % Identity: 100
- Alignment Length: 112
- Location of Alignment in SEQ ID NO 172: from 1 to

20 112

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 173
- Ceres seq\_id 1378583
- Location of start within SEQ ID NO 171: at 256 nt.

25

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- 30 - DnaJ domain
- Location within SEQ ID NO 173: from 9 to 53 aa.

30

## (Dp) Related Amino Acid Sequences

- 35 - Alignment No. 112
- gi No. 4263775
- Description:
- % Identity: 100
- Alignment Length: 112
- Location of Alignment in SEQ ID NO 173: from 1 to

35

40 64

## Maximum Length Sequence:

## related to:

## Clone IDs:

13399

## (Ac) cDNA Polynucleotide Sequence

- 45 - Pat. Appln. SEQ ID NO 174
- Ceres seq\_id 1383462
- Alternative transcription start site(s) located in SEQ ID NO 174:

50

-2,2,3,4,5,6,8,14,18,24

165

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 175
- Ceres seq\_id 1383463
- Location of start within SEQ ID NO 174: at 3 nt.

5

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Plant lipid transfer protein family
- Location within SEQ ID NO 175: from 46 to 119 aa.

10

## (Dp) Related Amino Acid Sequences

- Alignment No. 113
- gi No. 3128176
- Description:
- % Identity: 37
- Alignment Length: 194
- Location of Alignment in SEQ ID NO 175: from 23

to 197

20

## (D) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 176
- Ceres seq\_id 1383464
- Location of start within SEQ ID NO 174: at 42 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 176: at 24 aa.

25

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Plant lipid transfer protein family
- Location within SEQ ID NO 176: from 33 to 106 aa.

30

## (Dp) Related Amino Acid Sequences

- Alignment No. 114
- gi No. 3128176
- Description:
- % Identity: 37
- Alignment Length: 194
- Location of Alignment in SEQ ID NO 176: from 10

40 to 184

## (3) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 177
- Ceres seq\_id 1383465
- Location of start within SEQ ID NO 174: at 90 nt.

45

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Plant lipid transfer protein family
- Location within SEQ ID NO 177: from 17 to 90 aa.

50

166

## (Dp) Related Amino Acid Sequences

- Alignment No. 115
- gi No. 3128176
- Description:
- % Identity: 37
- Alignment Length: 194
- Location of Alignment in SEQ ID NO 177: from 1 to

168

## (10) Maximum Length Sequence:

## related to:

## Clone IDs:

- 156375
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 178
- Ceres seq\_id 1386215
- Alternative transcription start site(s) located in SEQ ID NO 178:

-38,12,17,18,19,20,26

20

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 179
- Ceres seq\_id 1386216
- Location of start within SEQ ID NO 178: at 2 nt.

25

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Photosystem I psaG / psaK
- Location within SEQ ID NO 179: from 77 to 138 aa.

30

## (Dp) Related Amino Acid Sequences

- Alignment No. 116
- gi No. 3885511
- Description:
- % Identity: 79.3
- Alignment Length: 112
- Location of Alignment in SEQ ID NO 179: from 28

to 138

40

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 180
- Ceres seq\_id 1386217
- Location of start within SEQ ID NO 178: at 83 nt.

45

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Photosystem I psaG / psaK
- Location within SEQ ID NO 180: from 50 to 111 aa.

50

## (Dp) Related Amino Acid Sequences

- Alignment No. 117

167

- gi No. 3885511
- Description:
- % Identity: 79.3
- Alignment Length: 112
- Location of Alignment in SEQ ID NO 180: from 1 to

5

111

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 181
- Ceres seq\_id 1386218
- Location of start within SEQ ID NO 178: at 95 nt.

10

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Photosystem I psaG / psaK
- Location within SEQ ID NO 181: from 46 to 107 aa.

15

## (Dp) Related Amino Acid Sequences

- Alignment No. 118
- gi No. 3885511
- Description:
- % Identity: 79.3
- Alignment Length: 112
- Location of Alignment in SEQ ID NO 181: from 1 to

20

25 107

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 182
- Ceres seq\_id 202556
- Location of start within SEQ ID NO 178: at 502 nt.

30

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

Maximum Length Sequence:  
related to:

Clone IDs:

21233

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 183
- Ceres seq\_id 1388499
- Alternative transcription start site(s) located in SEQ ID NO 183:

2,5,6,10,16,30,346,349

45

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 184
- Ceres seq\_id 1388500
- Location of start within SEQ ID NO 183: at 62 nt.

50

168

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- (Dp) Related Amino Acid Sequences
- Alignment No. 119
- gi No. 2829899
- Description:
- % Identity: 49.3
- Alignment Length: 150
- Location of Alignment in SEQ ID NO 184: from 2 to

5

10 150

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 185
- Ceres seq\_id 1388501
- Location of start within SEQ ID NO 183: at 122 nt.

15

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- (Dp) Related Amino Acid Sequences
- Alignment No. 120
- gi No. 2829899
- Description:
- % Identity: 49.3
- Alignment Length: 150
- Location of Alignment in SEQ ID NO 185: from 1 to

20

25 130

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 186
- Ceres seq\_id 1388502
- Location of start within SEQ ID NO 183: at 266 nt.

30

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- (Dp) Related Amino Acid Sequences
- Alignment No. 121
- gi No. 2829899
- Description:
- % Identity: 49.3
- Alignment Length: 150
- Location of Alignment in SEQ ID NO 186: from 1 to

35

40 82

(Ba) Polypeptide Activities: Similar to major latex protein activities.

45

Maximum Length Sequence:  
related to:  
Clone IDs:

50

169

2:304

- (Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 187  
- Ceres seq\_id 1388519

5

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 188  
- Ceres seq\_id 1388520  
- Location of start within SEQ ID NO 187: at 1 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 188: at 38 aa.

10

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 122  
- gi No. 4584113  
- Description:  
- % Identity: 48.5  
- Alignment Length: 163  
- Location of Alignment in SEQ ID NO 188: from 20 to 182

20

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 189  
- Ceres seq\_id 1388521  
- Location of start within SEQ ID NO 187: at 58 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 189: at 19 aa.

30

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 123  
- gi No. 4584110  
- Description:  
- % Identity: 48.5  
- Alignment Length: 163  
- Location of Alignment in SEQ ID NO 189: from 1 to 63

40

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 190  
- Ceres seq\_id 1388522  
- Location of start within SEQ ID NO 187: at 73 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 190: at 14 aa.

43

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 125

- gi No. 3927834

- Description:

- % Identity: 31.5

- Alignment Length: 108

- Location of Alignment in SEQ ID NO 193: from 1 to 43

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 134

- Ceres seq\_id 1388566

170

(Dp) Related Amino Acid Sequences

- Alignment No. 124  
- gi No. 4584110  
- Description:  
- % Identity: 48.5  
- Alignment Length: 163  
- Location of Alignment in SEQ ID NO 190: from 1 to 158

158

- 10 (Ba) Polypeptide Activities: Similar to pollen specific protein activities.

Maximum Length Sequence:

related to:

Clone IDs:

2153

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 191  
- Ceres seq\_id 1388563  
- Alternative Transcription start site(s) located in SEQ ID NO 191:  
-3,-2,-1,13,15,146

20

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 192  
- Ceres seq\_id 1388564  
- Location of start within SEQ ID NO 191: at 2 nt.

25

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 193  
- Ceres seq\_id 1388565  
- Location of start within SEQ ID NO 191: at 306 nt.

35

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 125

- gi No. 3927834

- Description:

- % Identity: 31.5

- Alignment Length: 108

- Location of Alignment in SEQ ID NO 193: from 1 to 43

43

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 134

- Ceres seq\_id 1388566

50

171

- Location of start within SEQ ID NO 191: at 455 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 194: at 36 aa.

5 (Ba) Polypeptide Activities: Similar to hydroxproline-rich glycoprotein activities.

Maximum Length Sequence:

related to:  
Clone IDs:

22488

Public Genomic DNA:

gi No: 5708384

Predicted Exons:

INIT 82228 ... 82063 OCKHAMG-CDS  
TERM 81730 ... 81361 OCKHAMG-CDS

gi No: 5732090

Predicted Exons:

INIT 21361 ... 21136 OCKHAMG-CDS  
TERM 20863 ... 20454 OCKHAMG-CDS

gi No: 5870169

Predicted Exons:

INIT 89258 ... 89423 OCKHAMG-CDS  
TERM 89756 ... 90105 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 195  
- Ceres seq\_id 1388793

- Alternative transcription start site(s) located in SEQ

30 ID NO 195:

7,14,15,20,24,30,33,42,44,58,63,64,66,76,94,97,98,99,101,102  
103,104,105,106,108,109,110,115,116,117

35 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 196

- Ceres seq\_id 1388794

- Location of start within SEQ ID NO 195: at 154 nt.

- Location of Signal Peptide Cleavage Site within SEQ

40 ID NO 196: at 19 aa.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 126

- gi No. 4584110

- Description:

- % Identity: 48.5

- Alignment Length: 163

- Location of Alignment in SEQ ID NO 196: from 1 to

163

172

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 197

- Ceres seq\_id 1388795

- Location of start within SEQ ID NO 195: at 169 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 197: at 14 aa.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 127

- gi No. 4584110

- Description:

- % Identity: 48.5

- Alignment Length: 163

- Location of Alignment in SEQ ID NO 197: from 1 to

158

20 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 198

- Ceres seq\_id 1388796

- Location of start within SEQ ID NO 195: at 205 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 128

- gi No. 4584110

- Description:

- % Identity: 48.5

- Alignment Length: 163

- Location of Alignment in SEQ ID NO 198: from 1 to

146

35 (Ba) Polypeptide Activities: Similar to pollen specific

protein activities.

Maximum Length Sequence:

related to:

Clone IDs:

26569

Public Genomic DNA:

gi No: 6143856

Predicted Exons:

INTR 78628 ... 78255 OCKHAMG-CDNA

INTR 78160 ... 78004 OCKHAMG-CDNA

INTR 78624 ... 78255 OCKHAMG-CDNA

INTR 78160 ... 78004 OCKHAMG-CDNA

50

173  
 INTR 78624 ... 78255 OCKHANG-CDNA  
 INTR 78624 ... 78006 OCKHANG-CDNA  
 INTR 78624 ... 78283 OCKHANG-CDNA  
 gi No: 6223633  
 Predicted Exons:  
 INTR 78629 ... 78256 OCKHANG-CDNA  
 INTR 78161 ... 78005 OCKHANG-CDNA  
 INTR 78625 ... 78256 OCKHANG-CDNA  
 INTR 78161 ... 78005 OCKHANG-CDNA  
 INTR 78625 ... 78256 OCKHANG-CDNA  
 INTR 78161 ... 78007 OCKHANG-CDNA  
 INTR 78625 ... 78284 OCKHANG-CDNA  
 INTR 78587 ... 78256 OCKHANG-CDS  
 TERM 78161 ... 78152 OCKHANG-CDS  
 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 199  
 - Ceres seq\_id 1392041  
 - Alternative transcription start site(s) located in SEQ ID NO 199:  
 3,4,5,6,7,12  
 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 200  
 - Ceres seq\_id 1392042  
 - Location of start within SEQ ID NO 199: at 45 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 200: at 19 aa.  
 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
 - Plant lipid transfer protein family  
 - Location within SEQ ID NO 200: from 23 to 112 aa.  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 129  
 - gi No. 2497753  
 - Description:  
 - % Identity: 47  
 - Alignment Length: 117  
 - Location of Alignment in SEQ ID NO 200: from 4 to 113  
 Maximum Length Sequence:  
 related to:  
 Clone IDs:  
 28475

174  
 Public Genomic DNA:  
 gi No: 5731404  
 Predicted Exons:  
 INTR 70440 ... 70346 OCKHANG-CDNA  
 INTR 70076 ... 69865 OCKHANG-CDNA  
 INTR 69308 ... 69434 OCKHANG-CDNA  
 INTR 69358 ... 69284 OCKHANG-CDNA  
 INTR 69195 ... 69115 OCKHANG-CDNA  
 INTR 69033 ... 68959 OCKHANG-CDNA  
 INTR 68068 ... 68695 OCKHANG-CDNA  
 INTR 68581 ... 68344 OCKHANG-CDNA  
 INTR 70056 ... 69865 OCKHANG-CDS  
 INTR 69508 ... 69434 OCKHANG-CDS  
 INTR 69358 ... 69284 OCKHANG-CDS  
 INTR 69195 ... 69115 OCKHANG-CDS  
 INTR 69033 ... 68959 OCKHANG-CDS  
 INTR 68868 ... 68695 OCKHANG-CDS  
 TERM 68581 ... 68346 OCKHANG-CDS  
 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 201  
 - Ceres seq\_id 1393342  
 - Alternative transcription start site(s) located in SEQ ID NO 201:  
 16,48  
 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 202  
 - Ceres seq\_id 1393343  
 - Location of start within SEQ ID NO 201: at 118 nt.  
 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 130  
 - gi No. 3386621  
 - Description:  
 - % Identity: 83  
 - Alignment Length: 224  
 - Location of Alignment in SEQ ID NO 202: from 2 to 225  
 (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot specific gene, plant specific gene.  
 Maximum Length Sequence:  
 related to:  
 Clone IDs:  
 29453  
 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 203

- 175  
- Ceres seq\_id 1393556  
- Alternative transcription start site(s) located in SEQ ID NO 203:  
7,11,14,23,37,41,42,57,61
- 5
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 204  
- Ceres seq\_id 1393557  
- Location of start within SEQ ID NO 203: at 3 nt.
- 10
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 131  
- gi No. 1082054  
- Description:  
- % Identity: 36.1  
- Alignment Length: 61  
- Location of Alignment in SEQ ID NO 204: from 23 to 82
- 15
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 205  
- Ceres seq\_id 1393558  
- Location of start within SEQ ID NO 203: at 69 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 205: at 41 aa.
- 20
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 132  
- gi No. 1082054  
- Description:  
- % Identity: 36.1  
- Alignment Length: 61  
- Location of Alignment in SEQ ID NO 205: from 1 to 60
- 30
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 206  
- Ceres seq\_id 1393559  
- Location of start within SEQ ID NO 203: at 72 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 206: at 40 aa.
- 35
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 133  
- gi No. 1082054
- 40
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 207  
- Ceres seq\_id 1396784  
- Location of start within SEQ ID NO 207: at 33 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 207: at 19 aa.
- 45
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 135  
- gi No. 4512613  
- Description:  
- % Identity: 99.2
- 50

- 176  
- Description:  
- % Identity: 36.1  
- Alignment Length: 61  
- Location of Alignment in SEQ ID NO 206: from 1 to 59
- 5
- (Ba) Polypeptide Activities: Similar to transmembrane copper transporter protein activities.
- 10
- Maximum Length Sequence:  
related to:  
Clone IDs:  
42384  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 207  
- Ceres seq\_id 1396782  
- Alternative transcription start site(s) located in SEQ ID NO 207:  
-1,29
- 15
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 208  
- Ceres seq\_id 1396783  
- Location of start within SEQ ID NO 207: at 3 nt.
- 20
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 134  
- gi No. 4512613  
- Description:  
- % Identity: 99.2  
- Alignment Length: 122  
- Location of Alignment in SEQ ID NO 208: from 1 to 132
- 25
- (3) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 209  
- Ceres seq\_id 1396784  
- Location of start within SEQ ID NO 207: at 33 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 209: at 19 aa.
- 30
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 135  
- gi No. 4512613  
- Description:  
- % Identity: 99.2
- 35
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 210  
- Ceres seq\_id 1396785  
- Location of start within SEQ ID NO 207: at 33 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 210: at 19 aa.
- 40
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 136  
- gi No. 4512613  
- Description:  
- % Identity: 99.2
- 45
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 211  
- Ceres seq\_id 1396786  
- Location of start within SEQ ID NO 207: at 33 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 211: at 19 aa.
- 50

- 122
- 177
- Alignment Length: 122
  - Location of Alignment in SEQ ID NO 209: from 1 to
- 5 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 210
  - Ceres seq\_id 1396785
  - Location of start within SEQ ID NO 207: at 45 nt.
- 10 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 136
  - gi No. 4512613
  - Description:
  - % Identity: 99.2
  - Alignment Length: 122
  - Location of Alignment in SEQ ID NO 210: from 1 to
- 15
- 118
- (Ba) Polypeptide Activities: Similar to intergenic region of cyt2-MDH1 of yeast activities.
- 25 Maximum Length Sequence:
- related to:
- Clone IDs:
- 42402
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 211
  - Ceres seq\_id 1396802
  - Alternative transcription start site(s) located in SEQ ID NO 211: -145, -56
- 35 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 212
  - Ceres seq\_id 1396803
  - Location of start within SEQ ID NO 211: at 1 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 212: at 19 aa.
- 40
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 137
  - gi No. 2463339
  - Description:
  - % Identity: 38.6
  - Alignment Length: 70
  - Location of Alignment in SEQ ID NO 212: from 136 to 205

178

- 5 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 213
  - Ceres seq\_id 1396804
  - Location of start within SEQ ID NO 211: at 136 nt.
- 10 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 138
  - gi No. 2463339
  - Description:
  - % Identity: 38.6
  - Alignment Length: 70
  - Location of Alignment in SEQ ID NO 213: from 91 to 160
- 15
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 214
  - Ceres seq\_id 1396805
  - Location of start within SEQ ID NO 211: at 172 nt.
- 20
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 139
  - gi No. 2463339
  - Description:
  - % Identity: 38.6
  - Alignment Length: 70
  - Location of Alignment in SEQ ID NO 214: from 79 to 148
- 30
- (Ba) Polypeptide Activities: Similar to rbcX protein activities.
- 35 Maximum Length Sequence:
- related to:
- Clone IDs:
- 5105
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 215
  - Ceres seq\_id 1397130
  - Alternative transcription start site(s) located in SEQ ID NO 215: -1
- 45
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 216
  - Ceres seq\_id 1397131
  - Location of start within SEQ ID NO 215: at 1 nt.



179

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 140
- gi No. 2827551
- Description:
- % Identity: 63.1
- Alignment Length: 179
- Location of Alignment in SEQ ID NO 216: from 2 to

10 176

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 217
- Ceres seq\_id 1397132
- Location of start within SEQ ID NO 215: at 52 nt.

15

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 141
- gi No. 2827551
- Description:
- % Identity: 63.1
- Alignment Length: 179
- Location of Alignment in SEQ ID NO 217: from 1 to

25

159

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 218
- Ceres seq\_id 1397133
- Location of start within SEQ ID NO 215: at 211 nt.

30

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 142
- gi No. 2827551
- Description:
- % Identity: 63.1
- Alignment Length: 179
- Location of Alignment in SEQ ID NO 218: from 1 to

40

106

(Ba) Polypeptide Activities: Similar to meth CpG binding  
protein activities.

Maximum Length Sequence:

related to:  
Clone IDs:

50

180

92204

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 219
- Ceres seq\_id 1398004
- Alternative transcription start site(s) located in SEQ ID NO 219:
- 2,3,5,8,28

5

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 220
- Ceres seq\_id 1398005
- Location of start within SEQ ID NO 219: at 68 nt.

10

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 143
- gi No. 4454037
- Description:
- % Identity: 99.3
- Alignment Length: 151
- Location of Alignment in SEQ ID NO 220: from 1 to 151

15

(Ba) Polypeptide Activities: Similar to major latex protein  
activities.

20

25

Maximum Length Sequence:

related to:

Clone IDs:

98584

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 221
- Ceres seq\_id 1399370

30

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 222
- Ceres seq\_id 1399371
- Location of start within SEQ ID NO 221: at 3 nt.

35

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 144
- gi No. 4886285
- Description:
- % Identity: 28.8
- Alignment Length: 59
- Location of Alignment in SEQ ID NO 222: from 25

45

50 83

- 181
- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 223  
 - Ceres seq\_id 1399372  
 - Location of start within SEQ ID NO 221: at 45 nt.
- 5 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 145  
 - gi No. 4886285  
 - Description:  
 - % Identity: 28.0  
 - Alignment Length: 59  
 - Location of Alignment in SEQ ID NO 223: from 11  
 15 to 69
- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 224  
 - Ceres seq\_id 1399373  
 - Location of start within SEQ ID NO 221: at 60 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ  
 ID NO 224: at 25 aa.
- 20 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 146  
 - gi No. 4886285  
 - Description:  
 - % Identity: 28.8  
 - Alignment Length: 59  
 - Location of Alignment in SEQ ID NO 224: from 6 to  
 64
- 25 (Ba) Polypeptide Activities: Similar to outer envelope  
 membrane protein in chloroplast in pea activities.
- 30 Maximum Length Sequence:  
 related to:  
 Clone IDs:  
 107400  
 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 225  
 - Ceres seq\_id 1425147  
 - Alternative transcription start site(s) located in SEQ  
 ID NO 225:  
 20,31,49,58,80
- 35 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 226

- 182
- Ceres seq\_id 1425148  
 - Location of start within SEQ ID NO 225: at 85 nt.
- 5 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 147  
 - gi No. 3510256  
 - Description:  
 - % Identity: 37.8  
 - Alignment Length: 158  
 - Location of Alignment in SEQ ID NO 226: from 1 to  
 135
- 10 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 227  
 - Ceres seq\_id 1425149  
 - Location of start within SEQ ID NO 225: at 241 nt.
- 15 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 148  
 - gi No. 3510256  
 - Description:  
 - % Identity: 37.8  
 - Alignment Length: 158  
 - Location of Alignment in SEQ ID NO 227: from 1 to  
 103
- 20 (Ba) Polypeptide Activities: Arabidopsis specific gene,  
 dicot specific gene, plant specific gene.
- 25 Maximum Length Sequence:  
 related to:  
 Clone IDs:  
 11073  
 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 228  
 - Ceres seq\_id 1441102  
 - Alternative transcription start site(s) located in SEQ  
 ID NO 228:  
 2,3,4,5,8,10,16,22,51,69,388
- 30 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 229  
 - Ceres seq\_id 1441103  
 - Location of start within SEQ ID NO 228: at 1 nt.
- 35 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)

183

## (Dp) Related Amino Acid Sequences

- Alignment No. 149
- gi No. 4337175
- Description:
- % Identity: 100
- Alignment Length: 179
- Location of Alignment in SEQ ID NO 229: from 19 to 197

5

10

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 230
- Ceres seq\_id 1441104
- Location of start within SEQ ID NO 228: at 55 nt.

15

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 150
- gi No. 4337175
- Description:
- % Identity: 100
- Alignment Length: 179
- Location of Alignment in SEQ ID NO 230: from 1 to 179

25

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 231
- Ceres seq\_id 1441105
- Location of start within SEQ ID NO 228: at 94 nt.

30

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 151
- gi No. 4337175
- Description:
- % Identity: 100
- Alignment Length: 179
- Location of Alignment in SEQ ID NO 231: from 1 to 166

40

## (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot specific gene, plant specific gene.

## Maximum Length Sequence:

related to:

Clone IDs:

267227

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 232
- Ceres seq\_id 1447480

50

184

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 233
- Ceres seq\_id 1447481
- Location of start within SEQ ID NO 232: at 52 nt.

5

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 152
- gi No. 3510256
- Description:
- % Identity: 54
- Alignment Length: 50
- Location of Alignment in SEQ ID NO 233: from 77 to 126

15

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 234
- Ceres seq\_id 1447482
- Location of start within SEQ ID NO 232: at 200 nt.

20

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 153
- gi No. 3510256
- Description:
- % Identity: 54
- Alignment Length: 50
- Location of Alignment in SEQ ID NO 234: from 25 to 74

30

## (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot specific gene, plant specific gene.

## Maximum Length Sequence:

related to:

Clone IDs:

268712

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 235
- Ceres seq\_id 1447577

45

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 236
- Ceres seq\_id 1447578
- Location of start within SEQ ID NO 235: at 2 nt.

50

185

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 5 - Alignment No. 154
- gi No. 5080769
- Description: 100
- % Identity: 100
- Alignment Length: 17
- Location of Alignment in SEQ ID NO 236: from 16
- 10 to 32

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 237
- Ceres seq\_id 1447579
- Location of start within SEQ ID NO 235: at 165 nt.
- 15

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 20 - Alignment No. 155
- gi No. 5080769
- Description: 91.5
- % Identity: 91.5
- Alignment Length: 153
- Location of Alignment in SEQ ID NO 237: from 1 to
- 25 115

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 238
- Ceres seq\_id 1447580
- Location of start within SEQ ID NO 235: at 298 nt.
- 30

(Ba) Polypeptide Activities: Similar to hydroxyproline-rich glycoprotein activities.

## Maximum Length Sequence:

related to:

Clone IDs:

94821

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 239
- Ceres seq\_id 1447922
- Alternative transcription start site(s) located in SEQ
- 45 ID NO 239:
- 12,16

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 240
- Ceres seq\_id 1447923
- Location of start within SEQ ID NO 239: at 81 nt.
- 50

186

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 5 - Alignment No. 156
- gi No. 3510256
- Description: 37.8
- % Identity: 37.8
- Alignment Length: 158
- Location of Alignment in SEQ ID NO 240: from 1 to
- 10 155

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 241
- Ceres seq\_id 1447924
- Location of start within SEQ ID NO 239: at 237 nt.
- 15

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 20 - Alignment No. 137
- gi No. 3510256
- Description: 37.8
- % Identity: 37.8
- Alignment Length: 158
- Location of Alignment in SEQ ID NO 241: from 1 to
- 25 103

(Ba) Polypeptide Activities: Arabidopsis specific gene, dict specific gene, plant specific gene.

## Maximum Length Sequence:

related to:

Clone IDs:

20539

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 242
- Ceres seq\_id 1448012
- Alternative transcription start site(s) located in SEQ
- 40 ID NO 242:
- 30,-6

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 243
- Ceres seq\_id 1448013
- Location of start within SEQ ID NO 242: at 79 nt.
- 45

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- 50 - Ribosomal protein S16
- Location within SEQ ID NO 243: from 9 to 70 aa.

187

## (Dp) Related Amino Acid Sequences

- Alignment No. 158
- gi No. 3096931
- Description:
- % Identity: 57.5
- Alignment Length: 113
- Location of Alignment in SEQ ID NO 243: from 1 to

112

10

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 244
- Ceres seq\_id 1448014
- Location of start within SEQ ID NO 242: at 139 nt.

15

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein S16
- Location within SEQ ID NO 244: from 1 to 53 aa.

20

## (Dp) Related Amino Acid Sequences

- Alignment No. 159
- gi No. 3096931
- Description:
- % Identity: 57.5
- Alignment Length: 113
- Location of Alignment in SEQ ID NO 244: from 1 to

92

30

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 245
- Ceres seq\_id 1448015
- Location of start within SEQ ID NO 242: at 232 nt.

35

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 160
- gi No. 3096931
- Description:
- % Identity: 57.5
- Alignment Length: 113
- Location of Alignment in SEQ ID NO 245: from 1 to

61

45

## Maximum Length Sequence:

related to:

Clone IDs:

34091

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 246

50

188

## (Dp) Related Amino Acid Sequences

- Ceres seq\_id 1448135
- Alternative transcription start site(s) located in SEQ ID NO 246:

-1

5

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 247
- Ceres seq\_id 1448136
- Location of start within SEQ ID NO 246: at 3 nt.

10

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 248
- Ceres seq\_id 1448137
- Location of start within SEQ ID NO 246: at 68 nt.

15

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein S14p/S29e
- Location within SEQ ID NO 248: from 3 to 54 aa.

20

## (Dp) Related Amino Acid Sequences

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 249
- Ceres seq\_id 1448138
- Location of start within SEQ ID NO 246: at 85 nt.

25

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 249
- Ceres seq\_id 1448138
- Location of start within SEQ ID NO 246: at 85 nt.

30

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 249
- Ceres seq\_id 1448138
- Location of start within SEQ ID NO 246: at 85 nt.

35

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 249
- Ceres seq\_id 1448138
- Location of start within SEQ ID NO 246: at 85 nt.

40

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

## Maximum Length Sequence:

related to:

Clone IDs:

39285

Public Genomic DNA:

gi No: 4406776

Predicted Exons:

SINGLE 61584 ... 61895 GENBANK

50

139

SINGLE 61584 ... 61895 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 250

- Ceres seq\_id 1448185

5 ID NO 250: Alternative transcription start site(s) located in SEQ

-36,-18,-17,16,19

(B) Polypeptide Sequence

10 - Pat. Appln. SEQ ID NO 251

- Ceres seq\_id 1448186

- Location of start within SEQ ID NO 250: at 41 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 162

- gi No. 4406787

- Description:

- % Identity: 100

- Alignment Length: 103

- Location of Alignment in SEQ ID NO 251: from 1 to

103

(B) Polypeptide Sequence

25 - Pat. Appln. SEQ ID NO 252

- Ceres seq\_id 1448187

- Location of start within SEQ ID NO 250: at 68 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 163

- gi No. 4406787

- Description:

- % Identity: 100

- Alignment Length: 103

- Location of Alignment in SEQ ID NO 252: from 1 to

94

(B) Polypeptide Sequence

40 - Pat. Appln. SEQ ID NO 253

- Ceres seq\_id 1448188

- Location of start within SEQ ID NO 250: at 89 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 164

- gi No. 4406787

- Description:

50

190

- % Identity: 100

- Alignment Length: 103

- Location of Alignment in SEQ ID NO 253: from 1 to 87

5 (Ba) Polypeptide Activities: Similar to NADH dehydrogenase protein activities

10 Maximum Length Sequence:

related to:

Clone IDs:

99119

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 254

- Ceres seq\_id 1450875

(B) Polypeptide Sequence

20 - Pat. Appln. SEQ ID NO 255

- Ceres seq\_id 1450876

- Location of start within SEQ ID NO 254: at 2 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 255: at 26 aa.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- ATPases associated with various cellular

activities (AAA)

30 - Location within SEQ ID NO 255: from 391 to 598

aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 165

- gi No. 4309734

- Description:

- % Identity: 86

- Alignment Length: 633

- Location of Alignment in SEQ ID NO 255: from 9 to

40 634

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 256

- Ceres seq\_id 1450877

- Location of start within SEQ ID NO 254: at 14 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 256: at 22 aa.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)



193

## (Dp) Related Amino Acid Sequences

- Alignment No. 170
- gi No. 4758714
- Description:
- % Identity: 41.9
- Alignment Length: 131
- Location of Alignment in SEQ ID NO 262: from 8 to

5

136

10

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 263
- Ceres seq\_id 1461850
- Location of start within SEQ ID NO 261: at 197 nt.

15

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- MAPEG family
- Location within SEQ ID NO 263: from 31 to 108 aa.

20

## (Dp) Related Amino Acid Sequences

- Alignment No. 171
- gi No. 4758714
- Description:
- % Identity: 41.9
- Alignment Length: 131
- Location of Alignment in SEQ ID NO 263: from 1 to

25

108

30

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 264
- Ceres seq\_id 1461851
- Location of start within SEQ ID NO 261: at 329 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 264: at 22 aa.

35

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- MAPEG family
- Location within SEQ ID NO 264: from 1 to 64 aa.

40

## (Dp) Related Amino Acid Sequences

- Alignment No. 172
- gi No. 4758714
- Description:
- % Identity: 41.9
- Alignment Length: 131
- Location of Alignment in SEQ ID NO 264: from 1 to

45

64

50 Maximum Length Sequence:  
related to:

194

## Clone IDs:

25093

- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 265
- Ceres seq\_id 1472772
- Alternative transcription start site(s) located in SEQ

5

ID NO 265:

-1,3

10

## (3) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 266
- Ceres seq\_id 1472773
- Location of start within SEQ ID NO 265: at 3 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 266: at 32 aa.

15

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Plant lipid transfer protein family
- Location within SEQ ID NO 266: from 45 to 108 aa.

20

## (Dp) Related Amino Acid Sequences

- Alignment No. 173
- gi No. 3062791
- Description:
- % Identity: 72.2
- Alignment Length: 90
- Location of Alignment in SEQ ID NO 266: from 21 to 110

25

to 110

30

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 267
- Ceres seq\_id 1472774
- Location of start within SEQ ID NO 265: at 27 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 267: at 24 aa.

35

ID NO 267: at 24 aa.

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Plant lipid transfer protein family
- Location within SEQ ID NO 267: from 37 to 100 aa.

40

## (Dp) Related Amino Acid Sequences

- Alignment No. 174
- gi No. 3062791
- Description:
- % Identity: 72.2
- Alignment Length: 90
- Location of Alignment in SEQ ID NO 267: from 13

45

50 to 102



195

## Maximum Length Sequence:

related to:

Clone IDs:

42300

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 268

- Ceres seq\_id 1533352

- Alternative transcription start site(s) located in SEQ

ID NO 268:

17

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 269

- Ceres seq\_id 1533353

- Location of start within SEQ ID NO 268: at 1 nt.

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- haloacid dehalogenase-like hydrolase

- Location within SEQ ID NO 269: from 82 to 266 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 175

- gi No. 3913203

- Description:

- % Identity: 32.8

- Alignment Length: 196

- Location of Alignment in SEQ ID NO 269: from 81

to 266

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 270

- Ceres seq\_id 1533354

- Location of start within SEQ ID NO 268: at 25 nt.

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- haloacid dehalogenase-like hydrolase

- Location within SEQ ID NO 270: from 74 to 258 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 176

- gi No. 3913203

- Description:

- % Identity: 32.8

- Alignment Length: 196

- Location of Alignment in SEQ ID NO 270: from 73

to 258

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 271

196

- Ceres seq\_id 1533355

- Location of start within SEQ ID NO 268: at 265 nt.

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- haloacid dehalogenase-like hydrolase

- Location within SEQ ID NO 271: from 1 to 178 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 177

- gi No. 3913203

- Description:

- % Identity: 32.8

- Alignment Length: 196

- Location of Alignment in SEQ ID NO 271: from 1 to

178

Maximum Length Sequence:

related to:

Clone IDs:

158412

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 272

- Ceres seq\_id 1534544

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 273

- Ceres seq\_id 1534545

- Location of start within SEQ ID NO 272: at 2 nt.

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 178

- gi No. 2317676

- Description:

- % Identity: 37.2

- Alignment Length: 115

- Location of Alignment in SEQ ID NO 273: from 22

to 134

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 274

- Ceres seq\_id 1534546

- Location of start within SEQ ID NO 272: at 524 nt.

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 179

197

- gi No. 4836939
- Description:
- % Identity: 47.3
- Alignment Length: 150
- Location of Alignment in SEQ ID NO 274: from 19 to 166

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 275
- Ceres seq\_id 1534547
- Location of start within SEQ ID NO 272: at 815 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 180
- gi No. 4836939
- Description:
- % Identity: 47.3
- Alignment Length: 150
- Location of Alignment in SEQ ID NO 275: from 1 to 69

(Ba) Polypeptide Activities: Similar to calcium independent phospholipase A2 protein activities, and acy-protein thioesterase activities.

## Maximum Length Sequence:

related to:

Clone IDs:

108109

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 276
- Ceres seq\_id 1567172
- Alternative transcription start site(s) located in SEQ ID NO 276: 22

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 277
- Ceres seq\_id 1567173
- Location of start within SEQ ID NO 276: at 152 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 181
- gi No. 1076331

198

- Description: CONSTANS protein - Arabidopsis thaliana >gi1161514|emb|CAA64407| (X94937) CONSTANS protein [Arabidopsis thaliana]
- % Identity: 49.2
- Alignment Length: 63
- Location of Alignment in SEQ ID NO 277: from 12 to 74

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 278
- Ceres seq\_id 1567174
- Location of start within SEQ ID NO 276: at 155 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 182
- gi No. 1076301
- Description: CONSTANS protein - Arabidopsis thaliana >gi1161514|emb|CAA64407| (X94937) CONSTANS protein [Arabidopsis thaliana]
- % Identity: 49.2
- Alignment Length: 63
- Location of Alignment in SEQ ID NO 278: from 11 to 73

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 279
- Ceres seq\_id 1567175
- Location of start within SEQ ID NO 276: at 443 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 183
- gi No. 1076301
- Description: CONSTANS protein - Arabidopsis thaliana >gi1161514|emb|CAA64407| (X94937) CONSTANS protein [Arabidopsis thaliana]
- % Identity: 67.4
- Alignment Length: 43
- Location of Alignment in SEQ ID NO 279: from 261 to 303

## Maximum Length Sequence:

related to:

Clone IDs:

168

Public Genomic DNA:

gi No: 4895213

Predicted Exons:

- 199  
INIT 22484 ... 22358 OCKHAMG-CDS  
TERM 21760 ... 21704 OCKHAMG-CDS  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 280  
- Ceres seq\_id 1567535  
- Alternative transcription start site(s) located in SEQ ID NO 280: 12,22,23,25,28,36,68
- 5
- 10 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 281  
- Ceres seq\_id 1567536  
- Location of start within SEQ ID NO 280: at 104 nt.
- 15 (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 184  
- gi No. 1351365  
- Description: UBIQUINCL-CYTOCHROME C REDUCTASE  
COMPLEX 6.7 KD PROTEIN (CR6) >gi12130002|pir||1568969  
ubiquincl--cytochrome-c reductase (EC 1.10.2.2) - potato  
>gi1633683|emb|CRA57768| (X82325) cytochrome c reductase  
subunit [Solanum tuberosum]  
- % Identity: 52.6  
- Alignment Length: 58  
- Location of Alignment in SEQ ID NO 281: from 2 to 57
- 20
- 25 Maximum Length Sequence:  
related to:  
Clone IDs:  
21305  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 282  
- Ceres seq\_id 1569689  
- Alternative transcription start site(s) located in SEQ ID NO 282: 25,29,49
- 30
- 35 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 283  
- Ceres seq\_id 1569690  
- Location of start within SEQ ID NO 282: at 1 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 283: at 32 aa.
- 40
- 45 (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences
- 50

- 200  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 284  
- Ceres seq\_id 1569691  
- Location of start within SEQ ID NO 282: at 92 nt.
- 5
- 10 (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 185  
- gi No. 2191138  
- Description: (AF007269) A\_IG002N01.18 gene  
product [Arabidopsis thaliana]  
- % Identity: 19.5  
- Alignment Length: 87  
- Location of Alignment in SEQ ID NO 284: from 3 to 83
- 15
- 20 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 285  
- Ceres seq\_id 1569692  
- Location of start within SEQ ID NO 282: at 248 nt.
- 25
- 30 (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 186  
- gi No. 2191138  
- Description: (AF007269) A\_IG002N01.18 gene  
product [Arabidopsis thaliana]  
- % Identity: 19.5  
- Alignment Length: 87  
- Location of Alignment in SEQ ID NO 285: from 1 to 31
- 35
- 40 (Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot specific gene, plant specific gene.  
Maximum Length Sequence:  
related to:  
Clone IDs:  
41988  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 286  
- Ceres seq\_id 1571042  
- Alternative transcription start site(s) located in SEQ ID NO 286: -1,2,3,4,7
- 45
- 50 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 287  
- Ceres seq\_id 1571043

201

- Location of start within SEQ ID NO 286: at 63 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 187
- gi No. 2129641
- Description: major latex protein type 1 -
- Arabidopsis thaliana >gi1107493|emb|CAA63026| (X91960) major
- latex protein type1 [Arabidopsis thaliana]
- % Identity: 71.3
- Alignment Length: 150
- Location of Alignment in SEQ ID NO 287: from 1 to

150

## Maximum Length Sequence:

related to:

Clone IDs:

16143

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 288
- Ceres seq\_id 1571079
- Alternative transcription start site(s) located in SEQ
- ID NO 288:
- 5,37,86,341,350

25

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 289
- Ceres seq\_id 1571080
- Location of start within SEQ ID NO 288: at 51 nt.

30

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Pathogenesis-related protein Bet v I family
- Location within SEQ ID NO 289: from 5 to 155 aa.

35

## (Dp) Related Amino Acid Sequences

- Alignment No. 188
- gi No. 1321731
- Description: (272439) major allergen Cor a 1
- [Corylus avellana]
- % Identity: 35.7
- Alignment Length: 159
- Location of Alignment in SEQ ID NO 289: from 5 to

155

## Maximum Length Sequence:

related to:

Clone IDs:

42101

## (Ac) cDNA Polynucleotide Sequence

50

202

- Pat. Appln. SEQ ID NO 290
- Ceres seq\_id 1572097

5

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 291
- Ceres seq\_id 1572098
- Location of start within: SEQ ID NO 290: at 63 nt.

10

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Adhesion lipoprotein
- Location within SEQ ID NO 291: from 18 to 191

aa.

15

## (Dp) Related Amino Acid Sequences

- Alignment No. 189
- gi No. 2129641
- Description: major latex protein type 1 -
- Arabidopsis thaliana >gi1107493|emb|CAA63026| (X91960) major
- latex protein type1 [Arabidopsis thaliana]
- % Identity: 69.5
- Alignment Length: 154
- Location of Alignment in SEQ ID NO 291: from 1 to

25 154

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 292
- Ceres seq\_id 1572099
- Location of start within SEQ ID NO 290: at 336 nt.

30

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Adhesion lipoprotein
- Location within SEQ ID NO 292: from 27 to 100 aa.

35

## (Dp) Related Amino Acid Sequences

- Alignment No. 190
- gi No. 2129641
- Description: major latex protein type 1 -
- Arabidopsis thaliana >gi1107493|emb|CAA63026| (X91960) major
- latex protein type1 [Arabidopsis thaliana]
- % Identity: 69.5
- Alignment Length: 154
- Location of Alignment in SEQ ID NO 292: from 1 to

63

## Maximum Length Sequence:

related to:

Clone IDs:

27643

50

203

- (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 293  
 - Ceres seq\_id 1572890  
 - Alternative transcription start site(s) located in SEQ ID NO 293:  
 -10

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 294  
 - Ceres seq\_id 1572891  
 - Location of start within SEQ ID NO 293: at 1 nt.

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Photosystem I psaG / psaK  
 - Location within SEQ ID NO 294: from 71 to 151 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 191  
 - gi No. 3885511  
 - Description: (AF084200) similar to PSI-K subunit of photosystem I from barley (Medicago sativa)  
 - % Identity: 80.3  
 - Alignment Length: 128  
 - Location of Alignment in SEQ ID NO 294: from 22 to 148

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 295  
 - Ceres seq\_id 1572892  
 - Location of start within SEQ ID NO 293: at 64 nt.

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Photosystem I psaG / psaK  
 - Location within SEQ ID NO 295: from 50 to 130 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 192  
 - gi No. 3885511  
 - Description: (AF084200) similar to PSI-K subunit of photosystem I from barley (Medicago sativa)  
 - % Identity: 80.3  
 - Alignment Length: 128  
 - Location of Alignment in SEQ ID NO 295: from 1 to 127

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 296  
 - Ceres seq\_id 1572893  
 - Location of start within SEQ ID NO 293: at 76 nt.

204

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Photosystem I psaG / psaK  
 - Location within SEQ ID NO 296: from 46 to 126 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 193  
 - gi No. 3885511  
 - Description: (AF084200) similar to PSI-K subunit of photosystem I from barley (Medicago sativa)  
 - % Identity: 80.3  
 - Alignment Length: 128  
 - Location of Alignment in SEQ ID NO 296: from 1 to 123

## Maximum Length Sequence:

related to:

Clone IDs:

- 33027  
 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 297  
 - Ceres seq\_id 1573606  
 - Alternative transcription start site(s) located in SEQ ID NO 297:  
 7, 8, 13

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 298  
 - Ceres seq\_id 1573607  
 - Location of start within SEQ ID NO 297: at 56 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 298: at 19 aa.

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Pollen proteins Ole e I family  
 - Location within SEQ ID NO 298: from 32 to 131 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 194  
 - gi No. 4584110  
 - Description: (AJ133639) SHA7 protein (Arabidopsis thaliana)  
 - % Identity: 48.5  
 - Alignment Length: 163  
 - Location of Alignment in SEQ ID NO 298: from 1 to 163

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 299

205

- Ceres seq\_id 1573608
- Location of start within SEQ ID NO 297: at 71 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 299: at 14 aa.

5

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Pollen proteins Ole e I family
- Location within SEQ ID NO 299: from 27 to 126 aa.

10

(Dp) Related Amino Acid Sequences

- Alignment No. 195
- gi No. 4584110
- Description: (AJ133639) SAH7 protein [Arabidopsis thaliana]

15

- % Identity: 48.5

- Alignment Length: 163

- Location of Alignment in SEQ ID NO 299: from 1 to

20

138

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 300
- Ceres seq\_id 1573609
- Location of start within SEQ ID NO 297: at 107 nt.

25

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Pollen proteins Ole e I family
- Location within SEQ ID NO 300: from 15 to 114 aa.

30

(Dp) Related Amino Acid Sequences

- Alignment No. 196
- gi No. 4584110
- Description: (AJ133639) SAH7 protein [Arabidopsis thaliana]

35

- % Identity: 48.5

- Alignment Length: 163

- Location of Alignment in SEQ ID NO 300: from 1 to

146

Maximum Length Sequence:  
related to:

Clone IDs:

31422

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 301
- Ceres seq\_id 1573861

45

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 302

50

206

- Ceres seq\_id 1573862
- Location of start within SEQ ID NO 301: at 67 nt.

5

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 197
- gi No. 2191138
- Description: (AF007269) A\_IG02N01.18 gene product [Arabidopsis thaliana]

10

- % Identity: 28.1

- Alignment Length: 160

- Location of Alignment in SEQ ID NO 302: from 3 to

156

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 303
- Ceres seq\_id 1573863
- Location of start within SEQ ID NO 301: at 223 nt.

20

104

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 198
- gi No. 2191138
- Description: (AF007269) A\_IG02N01.18 gene product [Arabidopsis thaliana]

25

- % Identity: 28.1

- Alignment Length: 160

- Location of Alignment in SEQ ID NO 303: from 1 to

30

104

(Ba) Polypeptide Activities: Arabidopsis specific gene, dicot specific gene, plant specific gene.

35

Maximum Length Sequence:  
related to:

Clone IDs:

40916

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 304
- Ceres seq\_id 1574093

40

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 305
- Ceres seq\_id 1574094
- Location of start within SEQ ID NO 304: at 47 nt.

45

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Pat. Appln. SEQ ID NO 306

50

207

- Zinc finger, C3HC4 type (RING finger)
- Location within SEQ ID NO 305: from 132 to 192

aa.

- 5 (Dp) Related Amino Acid Sequences
- Alignment No. 199
  - gi No. 3790593
  - Description: (AF079185) RING-H2 finger protein RHY1a (Arabidopsis thaliana)
  - % Identity: 55.1
  - Alignment Length: 49
  - Location of Alignment in SEQ ID NO 305: from 148 to 196

- 15 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 306
  - Ceres seq\_id 1574095
  - Location of start within SEQ ID NO 304: at 56 nt.

- 20 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- Zinc finger, C3HC4 type (RING finger)
  - Location within SEQ ID NO 306: from 149 to 189

aa.

- 25 (Dp) Related Amino Acid Sequences
- Alignment No. 200
  - gi No. 3790593
  - Description: (AF079185) RING-H2 finger protein RHY1a (Arabidopsis thaliana)
  - % Identity: 55.1
  - Alignment Length: 49
  - Location of Alignment in SEQ ID NO 306: from 145 to 193

Maximum Length Sequence:

related to:

Clone IDs:

35447

40 Public Genomic DNA:

gi No: 4263694

Predicted Exons:

INIT 79640 ... 79806 GENBANK

TERM 79888 ... 80164 GENBANK

INTR 79508 ... 79806 OCKHAMG-CDNA

INTR 79888 ... 80309 OCKHAMG-CDNA

INIT 79640 ... 79806 OCKHAMG-CDS

TERM 79888 ... 80164 OCKHAMG-CDS

50 (Ac) cDNA Polynucleotide Sequence

208

- Pat. Applr. SEQ ID NO 307
- Ceres seq\_id 1580388
- Alternative transcription start site(s) located in SEQ ID NO 307:

- 5 2,3,4,5,6,7,63,94

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 308
  - Ceres seq\_id 1580389
  - Location of start within SEQ ID NO 307: at 133 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 308: at 27 aa.

- 15 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- Cystatin domain
  - Location within SEQ ID NO 308: from 87 to 141 aa.

(Dp) Related Amino Acid Sequences

- 20 - Alignment No. 201
- gi No. 2204077
  - Description: (D85823) extracellular insoluble cystatin [Daucus carota]
  - % Identity: 44.9

- 25 - Alignment Length: 136
- Location of Alignment in SEQ ID NO 308: from 6 to 141

(B) Polypeptide Sequence

- 30 - Pat. Appln. SEQ ID NO 309
- Ceres seq\_id 1580390
  - Location of start within SEQ ID NO 307: at 142 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 309: at 24 aa.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Cystatin domain

- 35 - Location within SEQ ID NO 309: from 84 to 138 aa.

(Dp) Related Amino Acid Sequences

- 40 - Alignment No. 202
- gi No. 2204077
  - Description: (D85823) extracellular insoluble cystatin [Daucus carota]
  - % Identity: 44.9
  - Alignment Length: 136
  - Location of Alignment in SEQ ID NO 309: from 3 to 138

Maximum Length Sequence:

209

related to:

Clone IDs:

22677

Public Genomic DNA:

gi No: 4972043

Predicted Exons:

SINGLE 18935 ... 19282 GENBANK

5

INTR 18856 ... 19527 OCKHAMG-CDNA

10

SINGLE 18935 ... 19282 OCKHAMG-CDS

gi No: 4972077

Predicted Exons:

INTR 60360 ... 61031 OCKHAMG-CDNA

15

SINGLE 60439 ... 60786 OCKHAMG-CDS

gi No: 5748495

Predicted Exons:

INTR 60360 ... 61031 OCKHAMG-CDNA

20

SINGLE 60439 ... 60786 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 310

- Ceres seq\_id 1582959

25 - Alternative transcription start site(s) located in SEQ

ID NO 310:

2,5,6,7,58,61

(B) Polypeptide Sequence

30 - Pat. Appln. SEQ ID NO 311

- Ceres seq\_id 1582960

- Location of start within SEQ ID NO 310: at 80 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 311: at 19 aa.

35

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Dehydrins

- Location within SEQ ID NO 311: from 37 to 113 aa.

40

(Dp) Related Amino Acid Sequences

- Alignment No. 203

- gi No. 4972049

- Description: (AL078470) glycine-rich protein like

- % Identity: 100

- Alignment Length: 115

- Location of Alignment in SEQ ID NO 311: from 1 to

115

50 Maximum Length Sequence:

210

related to:

Clone IDs:

15190

34118

Public Genomic DNA:

gi No: 4582444

Predicted Exons:

INIT 39128 ... 39157 GENBANK

INTR 39241 ... 39388 GENBANK

INTR 39469 ... 39640 GENBANK

INTR 39922 ... 40108 GENBANK

TERM 40198 ... 40413 GENBANK

INTR 39945 ... 40108 OCKHAMG-CDNA

INTR 40198 ... 40660 OCKHAMG-CDNA

INTR 39054 ... 39157 OCKHAMG-CDNA

INTR 39241 ... 39388 OCKHAMG-CDNA

INTR 39469 ... 39640 OCKHAMG-CDNA

INTR 39922 ... 40108 OCKHAMG-CDNA

INTR 40198 ... 40626 OCKHAMG-CDNA

INIT 39128 ... 39157 OCKHAMG-CDS

INTR 39241 ... 39388 OCKHAMG-CDS

INTR 39469 ... 39640 OCKHAMG-CDS

INTR 39922 ... 40108 OCKHAMG-CDS

TERM 40198 ... 40413 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 312

- Ceres seq\_id 1663221

30 - Alternative transcription start site(s) located in SEQ

ID NO 312:

-5,2,3,4,5,12,40,448

- Clone 15190 starts at 448 and ends at 1076 in cDNA.

- Clone 34118 starts at 1 and ends at 1042 in cDNA.

35

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 313

- Ceres seq\_id 1663222

- Location of start within SEQ ID NO 312: at 75 nt.

40

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Ribosomal protein S3, C-terminal domain.

- Location within SEQ ID NO 313: from 104 to 188

45

aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 204

- gi No. 133940

50



211

- Description: 40S RIBOSOMAL PROTEIN S3A (S1A)  
>gi170851|pir||R3XL3A ribosomal protein S3a - African clawed  
frog >gi165091|emb|CRA405921 (X57322); ribosomal protein S1a  
[Xenopus laevis]  
- % Identity: 80.1  
- Alignment Length: 231  
- Location of Alignment in SEQ ID NO 313: from 1 to 228

5

10

(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 314  
- Ceres seq\_id 1663223  
- Location of start within SEQ ID NO 312: at 204 nt.  
(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
- Ribosomal protein S3, C-terminal domain.  
- Location within SEQ ID NO 314: from 61 to 145 aa.

15

20

(Dp) Related Amino Acid Sequences  
- Alignment No. 205  
- gi No. 133940  
- Description: 40S RIBOSOMAL PROTEIN S3A (S1A)  
>gi170851|pir||R3XL3A ribosomal protein S3a - African clawed  
frog >gi165091|emb|CRA405921 (X57322); ribosomal protein S1a  
[Xenopus laevis]  
- % Identity: 80.1  
- Alignment Length: 231  
- Location of Alignment in SEQ ID NO 314: from 1 to 185

30

Maximum Length Sequence:  
related to:

35

Clone IDs:  
3996  
2749

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 315  
- Ceres seq\_id 1663275  
- Alternative transcription start site(s) located in SEQ ID NO 315:  
3,4,5,6,7,8,9,15  
- Clone 3996 starts at 1 and ends at 484 in cDNA.  
- Clone 2749 starts at 7 and ends at 488 in cDNA.

40

45

(3) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 316  
- Ceres seq\_id 1663276  
- Location of start within SEQ ID NO 315: at 75 nt.

50

212

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
- Ribosomal protein S14p/S29e  
- Location within SEQ ID NO 316: from 3 to 54 aa.

5

(Dp) Related Amino Acid Sequences

- Alignment No. 206  
- gi No. 4506717  
- Description: ref|NP\_001023.1|PRPS29| ribosomal  
protein S29 >gi1266972|sp|P30054|RS29 HUMAN 40S RIBOSOMAL  
PROTEIN S29 >gi1631884|pir||S30298 ribosomal protein S29 -  
rat >gi1362934|pir||S55919 ribosomal protein S29 - human  
protein S29 [Homo sapiens]  
- % Identity: 72.2  
- Alignment Length: 54  
- Location of Alignment in SEQ ID NO 316: from 1 to 54

10

15

20

(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 317  
- Ceres seq\_id 1663277  
- Location of start within SEQ ID NO 315: at 92 nt.

25

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences

Maximum Length Sequence:  
related to:

30

Clone IDs:  
114940  
25068

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 318  
- Ceres seq\_id 1665304  
- Alternative transcription start site(s) located in SEQ ID NO 318:  
-8,-3,-1,2,3,4,8,17,251  
- Clone 114940 starts at 1 and ends at 743 in cDNA.  
- Clone 25068 starts at 4 and ends at 760 in cDNA.

40

45

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 319  
- Ceres seq\_id 1665305  
- Location of start within SEQ ID NO 318: at 63 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences  
- Alignment No. 207  
- gi No. 2129641

50

213

- Description: major latex protein type 1 - Arabidopsis thaliana >gi1107493|emb|CAA63026| (X91960) major latex protein type 1 [Arabidopsis thaliana]
- % Identity: 71.3
- Alignment Length: 150
- Location of Alignment in SEQ ID NO 319: from 1 to 150

## Maximum Length Sequence:

related to:

Clone IDs:

94673

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 320

- Ceres seq\_id 1709970

- Alternative transcription start site(s) located in SEQ ID NO 320:

-2,-1,11,41,42,43,44,45,46,47,48,51

20 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 321

- Ceres seq\_id 1709971

- Location of start within SEQ ID NO 320: at 276 nt.

25 (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 208

- gi No. 2961300

- Description: (AJ225027) ribosomal protein L24

[Cicer arietinum]

- % Identity: 86.2

- Alignment Length: 160

- Location of Alignment in SEQ ID NO 321: from 1 to 35

107

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 322

- Ceres seq\_id 1709972

- Location of start within SEQ ID NO 320: at 424 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

Maximum Length Sequence:

related to:

Clone IDs:

112835

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 323

214

- Ceres seq\_id 1711273
- Alternative transcription start site(s) located in SEQ ID NO 323:
- 2,28,30,33,56,62,235

5

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 324

- Ceres seq\_id 1711274

- Location of start within SEQ ID NO 323: at 92 nt.

10

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Helix-turn-helix

- Location within SEQ ID NO 324: from 87 to 141 aa.

15

(Dp) Related Amino Acid Sequences

- Alignment No. 209

- gi No. 1632831

- Description: (Z49698) orf [Ricinus communis]

- % Identity: 81

- Alignment Length: 142

- Location of Alignment in SEQ ID NO 324: from 1 to 142

142

25 Maximum Length Sequence:

related to:

Clone IDs:

17878

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 325

- Ceres seq\_id 1715423

- Alternative transcription start site(s) located in SEQ ID NO 325:

2,4,7,66

35

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 326

- Ceres seq\_id 1715424

- Location of start within SEQ ID NO 325: at 25 nt.

40

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 210

- gi No. 4337175

- Description: (AC006416) ESTs gb|T20589,

gb|T04648, gb|AA597906, gb|T04111, gb|R84180, gb|R65428,

gb|T44439, gb|T76570, gb|R90004, gb|T45020, gb|T42457,

gb|T20921, gb|AA042762 and gb|AA720210 come from this gene.

[Arabidopsis thaliana]

- % Identity: 44.8

50

215

- Alignment Length: 67
- Location of Alignment in SEQ ID NO 326: from 15

to 81

5

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 327
  - Ceres seq\_id 1715425
  - Location of start within SEQ ID NO 325: at 40 nt.

10

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 211
  - gi No. 4337175

15

- Description: (AC006416) ESTs gb|T20589, gb|T04648, gb|AA597906, gb|T04111, gb|R84180, gb|R65428, gb|T44439, gb|T76570, gb|R90004, gb|T45020, gb|T42457, gb|T20921, gb|AA042762 and gb|AA720210 come from this gene. [Arabidopsis thaliana]
- % Identity: 44.8
- Alignment Length: 67
- Location of Alignment in SEQ ID NO 327: from 10

20

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 328
  - Ceres seq\_id 1715426
  - Location of start within SEQ ID NO 325: at 260 nt.

25

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 212
  - gi No. 4337175
  - Description: (AC006416) ESTs gb|T20589, gb|T04648, gb|AA597906, gb|T04111, gb|R84180, gb|R65428, gb|T44439, gb|T76570, gb|R90004, gb|T45020, gb|T42457, gb|T20921, gb|AA042762 and gb|AA720210 come from this gene. [Arabidopsis thaliana]
  - % Identity: 56.1
  - Alignment Length: 57
  - Location of Alignment in SEQ ID NO 328: from 5 to

30

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 331
  - Ceres seq\_id 1715964
  - Location of start within SEQ ID NO 329: at 114 nt.

35

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 214
  - gi No. 1710581
  - Description: 60S RIBOSOMAL PROTEIN L9
  - Ribosomal protein L6
  - Location within SEQ ID NC 331: from 3 to 185 aa.

40

- (Dp) Related Amino Acid Sequences
- Alignment No. 214
  - gi No. 1710581
  - Description: 60S RIBOSOMAL PROTEIN L9
  - Ribosomal protein L6
  - Location within SEQ ID NC 331: from 3 to 185 aa.

61

45

- (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot specific gene, plant specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

26744

50 185

216

- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 329
  - Ceres seq\_id 1715962
  - Alternative transcription start site(s) located in SEQ

5

- ID NO 329:
- 2, 7, 9, 10, 26, 28, 30, 31, 32, 38, 39, 42, 43, 47, 48, 194, 247

10

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 330
  - Ceres seq\_id 1715963
  - Location of start within SEQ ID NO 329: at 87 nt.

15

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- Ribosomal protein L6
  - Location within SEQ ID NO 330: from 12 to 194 aa.

20

- (Dp) Related Amino Acid Sequences
- Alignment No. 213
  - gi No. 1710581
  - Description: 60S RIBOSOMAL PROTEIN L9
  - Ribosomal protein L6
  - Location within SEQ ID NO 330: from 12 to 194 aa.

25

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 331
  - Ceres seq\_id 1715964
  - Location of start within SEQ ID NO 329: at 114 nt.

194

30

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 214
  - gi No. 1710581
  - Description: 60S RIBOSOMAL PROTEIN L9
  - Ribosomal protein L6
  - Location within SEQ ID NC 331: from 3 to 185 aa.

35

- (Dp) Related Amino Acid Sequences
- Alignment No. 214
  - gi No. 1710581
  - Description: 60S RIBOSOMAL PROTEIN L9
  - Ribosomal protein L6
  - Location within SEQ ID NC 331: from 3 to 185 aa.

40

- (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot specific gene, plant specific gene.

45

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 331
  - Ceres seq\_id 1715964
  - Location of start within SEQ ID NO 329: at 114 nt.

Maximum Length Sequence:

related to:

Clone IDs:

26744

50 185

217

## (D) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 332
- Ceres seq\_id 1715965
- Location of start within SEQ ID NO 329: at 369 nt.

5

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein L6
- Location within SEQ ID NO 332: from 1 to 100 aa.

10

## (Dp) Related Amino Acid Sequences

- Alignment No. 215
- gi No. 1710581
- Description: 60S RIBOSOMAL PROTEIN L9
- >gi|2129720|pir||S71255 ribosomal protein L9 - Arabidopsis thaliana >gi|1107489|emb|CAA63024| (X91958) 60S ribosomal protein L9 [Arabidopsis thaliana]
- % Identity: 93.3
- Alignment Length: 195
- Location of Alignment in SEQ ID NO 332: from 1 to 100

20

## Maximum Length Sequence:

## related to:

## Clone IDs:

41712

## Public Genomic DNA:

gi No: 4512656

## Predicted Exons:

- |    |      |           |        |              |
|----|------|-----------|--------|--------------|
| 30 | INIT | 106774... | 106633 | GENBANK      |
|    | INTR | 106277... | 106194 | GENBANK      |
|    | INTR | 105920... | 105938 | GENBANK      |
|    | TERM | 105749... | 105630 | GENBANK      |
| 35 | INTR | 106841... | 106633 | OCKHANG-CDNA |
|    | INTR | 106277... | 106194 | OCKHANG-CDNA |
|    | INTR | 105920... | 105838 | OCKHANG-CDNA |
|    | INTR | 105749... | 105453 | OCKHANG-CDNA |

40

- |      |           |        |             |
|------|-----------|--------|-------------|
| INIT | 106774... | 106633 | OCKHANG-CDS |
| INTR | 106277... | 106194 | OCKHANG-CDS |
| INTR | 105920... | 105838 | OCKHANG-CDS |
| TERM | 105749... | 105630 | OCKHANG-CDS |

gi No: 4559319

## Predicted Exons:

- |    |      |           |       |              |
|----|------|-----------|-------|--------------|
| 45 | INIT | 10855 ... | 10714 | GENBANK      |
|    | INTR | 10358 ... | 10275 | GENBANK      |
|    | INTR | 10001 ... | 9919  | GENBANK      |
|    | TERM | 9830 ...  | 9711  | GENBANK      |
| 50 | INTR | 10922 ... | 10714 | OCKHANG-CDNA |

218

- |      |           |       |              |
|------|-----------|-------|--------------|
| INTR | 10358 ... | 10275 | OCKHANG-CDNA |
| INTR | 10001 ... | 9919  | OCKHANG-CDNA |
| INTR | 9830 ...  | 9534  | OCKHANG-CDNA |

5

- |      |           |       |             |
|------|-----------|-------|-------------|
| INIT | 10855 ... | 10714 | OCKHANG-CDS |
| INTR | 10358 ... | 10275 | OCKHANG-CDS |
| INTR | 10001 ... | 9919  | OCKHANG-CDS |
| TERM | 9830 ...  | 9711  | OCKHANG-CDS |

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 333
- Ceres seq\_id 1308584
- Alternative transcription start site(s) located in SEQ ID NO 333:

34

15

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 334
- Ceres seq\_id 1808585
- Location of start within SEQ ID NO 333: at 1 nt.

20

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Helix-turn-helix
- Location within SEQ ID NO 334: from 110 to 164

25 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 216
- gi No. 1632831
- Description: (249698) orf (Ricinus communis)
- % Identity: 81
- Alignment Length: 142
- Location of Alignment in SEQ ID NO 334: from 24 to 165

35

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 335
- Ceres seq\_id 1008586
- Location of start within SEQ ID NO 333: at 70 nt.

40

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Helix-turn-helix
- Location within SEQ ID NO 335: from 87 to 141 aa.

45

## (Dp) Related Amino Acid Sequences

- Alignment No. 217
- gi No. 1632831
- Description: (249698) orf (Ricinus communis)
- % Identity: 81
- Alignment Length: 142

50

142 219  
- Location of Alignment in SEQ ID NO 335: from 1 to

## Maximum Length Sequence:

related to:

Clone IDs:

42211

Public Genomic DNA:

gi No: 491400

Predicted Exons:

INTR 90159 ... 89532 OCKHAMG-CDNA

SINGLE 90124 ... 89615 OCKHAMG-CDS

gi No: 4938473

Predicted Exons:

SINGLE 2368 ... 1859 GENBANK

INTR 2403 ... 1776 OCKHAMG-CDNA

SINGLE 2368 ... 1859 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 336

- Ceres seq\_id 1808591

- Alternative transcription start site(s) located in SEQ

25 ID NO 336:

7,10

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 337

- Ceres seq\_id 1808592

- Location of start within SEQ ID NO 336: at 2 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 337: at 36 aa.

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 218

- gi No. 4926823

- Description: (AC004135) T17H7.8 [Arabidopsis

thaliana]

- % Identity: 46.1

- Alignment Length: 178

- Location of Alignment in SEQ ID NO 337: from 13

45 to 177

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 338

- Ceres seq\_id 1808593

- Location of start within SEQ ID NO 336: at 38 nt.

50

220  
- Location of Signal Peptide Cleavage Site within SEQ  
ID NO 338: at 24 aa.

## (C) Nomination and Annotation of Domains within

5 Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 219

- gi No. 4926823

- Description: (AC004135) T17H7.8 [Arabidopsis

10 thaliana]

- % Identity: 46.1

- Alignment Length: 178

- Location of Alignment in SEQ ID NO 338: from 1 to

165

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 339

- Ceres seq\_id 1808594

- Location of start within SEQ ID NO 336: at 164 nt.

20

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 220

- gi No. 4926823

- Description: (AC004135) T17H7.8 [Arabidopsis

thaliana]

- % Identity: 46.1

- Alignment Length: 178

- Location of Alignment in SEQ ID NO 339: from 1 to

30 123

(Ba) Polypeptide Activities: Arabidopsis specific gene,

dicot specific gene, plant specific gene.

## Maximum Length Sequence:

related to:

Clone IDs:

22-7

40 (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 340

- Ceres seq\_id 1920563

- Alternative transcription start site(s) located in SEQ

ID NO 340:

6

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 341

- Ceres seq\_id 1920564

- Location of start within SEQ ID NO 340: at 1 nt.

50

221

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- Oxidoreductase FAD/NAD-binding domain
- Location within SEQ ID NO 341: from 45 to 167 aa.

5

(Dp) Related Amino Acid Sequences

- Alignment No. 221
  - gi No. 729477
  - Description: FERREDOXIN--NADP REDUCTASE PRECURSOR
- (FNR) >gi|320548|pir||A44974 ferredoxin--NADP+ reductase (EC 1.18.1.2) precursor - common ice plant >gi|167256 (M25528) ferredoxin-NADP+ reductase precursor (fnrA; EC 1.6.7.1) [Mesembryanthemum]

10

- % Identity: 89.8

- Alignment Length: 197
- Location of Alignment in SEQ ID NO 341: from 1 to

15

197

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 342
- Ceres seq\_id 1920565
- Location of start within SEQ ID NO 340: at 112 nt.

20

(C) Nomination and Annotation of Domains within

- Predicted Polypeptide(s)
- Oxidoreductase FAD/NAD-binding domain
- Location within SEQ ID NO 342: from 8 to 130 aa.

25

(Dp) Related Amino Acid Sequences

- Alignment No. 222
  - gi No. 729477
  - Description: FERREDOXIN--NADP REDUCTASE PRECURSOR
- (FNR) >gi|320548|pir||A44974 ferredoxin--NADP+ reductase (EC 1.18.1.2) precursor - common ice plant >gi|167256 (M25528) ferredoxin-NADP+ reductase precursor (fnrA; EC 1.6.7.1) [Mesembryanthemum]

35

- % Identity: 89.8

- Alignment Length: 197
- Location of Alignment in SEQ ID NO 342: from 1 to

40 160

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 343
- Ceres seq\_id 1920566
- Location of start within SEQ ID NO 340: at 118 nt.

45

(C) Nomination and Annotation of Domains within

- Predicted Polypeptide(s)
- Oxidoreductase FAD/NAD-binding domain
- Location within SEQ ID NO 343: from 6 to 128 aa.

50

222

(Dp) Related Amino Acid Sequences

- Alignment No. 223
  - gi No. 729477
  - Description: FERREDOXIN--NADP REDUCTASE PRECURSOR
- (FNR) >gi|320548|pir||A44974 ferredoxin--NADP+ reductase (EC 1.18.1.2) precursor - common ice plant >gi|167256 (M25528) ferredoxin-NADP+ reductase precursor (fnrA; EC 1.6.7.1) [Mesembryanthemum]

5

- % Identity: 89.8

- Alignment Length: 197
- Location of Alignment in SEQ ID NO 343: from 1 to

10

158

Maximum Length Sequence:

- related to:

Clone IDs:

13864

Public Genomic DNA:

- gi No: 4303719
- Predicted Exons:
- INIT 50987 ... 50773 GENBANK
- INTR 49622 ... 49352 GENBANK
- INTR 49257 ... 49162 GENBANK
- INTR 48896 ... 48786 GENBANK
- INTR 48686 ... 48377 GENBANK
- INTR 48133 ... 48043 GENBANK
- INTR 47924 ... 47591 GENBANK
- INTR 47475 ... 47327 GENBANK
- TERM 47084 ... 46550 GENBANK

20

25

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 344
- Ceres seq\_id 1974419
- Alternative transcription start site(s) located in SEQ

ID NO 344:

-5

35

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 345
- Ceres seq\_id 1974420
- Location of start within SEQ ID NO 344: at 81 nt.

40

(C) Nomination and Annotation of Domains within

- Predicted Polypeptide(s)
- ATPases associated with various cellular activities (AAA)
- Location within SEQ ID NO 345: from 454 to 664 aa.

45

(Dp) Related Amino Acid Sequences

- Alignment No. 224
- gi No. 4006905

50

- 223
- Description: (Z99708) ATPase-like protein [Arabidopsis thaliana]
  - % Identity: 84.3
  - Alignment Length: 633
  - Location of Alignment in SEQ ID NO 345: from 72 to 700

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 346
  - Ceres seq\_id 1974421
  - Location of start within SEQ ID NO 344: at 96 nt.

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- ATPases associated with various cellular activities (AAA)
  - Location within SEQ ID NO 346: from 449 to 659 aa.

- (Dp) Related Amino Acid Sequences
- Alignment No. 225
  - gi No. 4006905
  - Description: (Z99708) ATPase-like protein [Arabidopsis thaliana]
  - % Identity: 84.3
  - Alignment Length: 633
  - Location of Alignment in SEQ ID NO 346: from 67 to 695

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 347
  - Ceres seq\_id 1974422
  - Location of start within SEQ ID NO 344: at 141 nt.

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- ATPases associated with various cellular activities (AAA)
  - Location within SEQ ID NO 347: from 434 to 644 aa.

- (Dp) Related Amino Acid Sequences
- Alignment No. 226
  - gi No. 4006935
  - Description: (Z99708) ATPase-like protein [Arabidopsis thaliana]
  - % Identity: 84.3
  - Alignment Length: 633
  - Location of Alignment in SEQ ID NO 347: from 52 to 680

- 224
- Maximum Length Sequence:  
related to:  
Clone IDs:  
36333
- 5 (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 348
  - Ceres seq\_id 1975983

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 349
  - Ceres seq\_id 1975984
  - Location of start within SEQ ID NO 348: at 90 nt.

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- Related Amino Acid Sequences
  - Alignment No. 227
  - gi No. 3688432
  - Description: (AJ011705) 40S ribosomal protein S10 [Lumbricus rubellus]
  - % Identity: 55.3
  - Alignment Length: 164
  - Location of Alignment in SEQ ID NO 349: from 32 to 192

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 350
  - Ceres seq\_id 1975985
  - Location of start within SEQ ID NO 348: at 135 nt.

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- Related Amino Acid Sequences
  - Alignment No. 228
  - gi No. 3688432
  - Description: (AJ011705) 40S ribosomal protein S10 [Lumbricus rubellus]
  - % Identity: 55.3
  - Alignment Length: 164
  - Location of Alignment in SEQ ID NO 350: from 17 to 177

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 351
  - Ceres seq\_id 1975986
  - Location of start within SEQ ID NO 348: at 183 nt.
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- Related Amino Acid Sequences

225

- Alignment No. 229
- gi No. 3688432
- Description: (AJ011705) 40S ribosomal protein S10 [Lumbricus rubellus]
- % Identity: 55.3
- Alignment Length: 164
- Location of Alignment in SEQ ID NO 351: from 1 to 161

5

- (Ba) Polypeptide Activities: Similar to glycine-rich RNA binding ABA inducible protein activities.

10

## Maximum Length Sequence:

related to:

Clone IDs:

24587

41072

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 352
- Ceres seq\_id 1976019
- Alternative transcription start site(s) located in SEQ ID NO 352:

13

ID NO 352:

13

- Clone 24587 starts at 13 and ends at in cDNA.

20

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 353
- Ceres seq\_id 1976020
- Location of start within SEQ ID NO 352: at 101 nt.

25

- (C) Nomination and Annotation of Domains within

(Dp) Related Amino Acid Sequences

- Alignment No. 230
- gi No. 3688432
- Description: (AJ011705) 40S ribosomal protein S10 [Lumbricus rubellus]
- % Identity: 55.3
- Alignment Length: 164
- Location of Alignment in SEQ ID NO 353: from 1 to

30

Predicted Polypeptide(s)

- Alignment No. 230
- gi No. 3688432
- Description: (AJ011705) 40S ribosomal protein S10 [Lumbricus rubellus]
- % Identity: 55.3
- Alignment Length: 164
- Location of Alignment in SEQ ID NO 353: from 1 to

35

(Lumbricus rubellus)

- Alignment No. 230

- gi No. 3688432

- Description: (AJ011705) 40S ribosomal protein S10 [Lumbricus rubellus]

- % Identity: 55.3

- Alignment Length: 164

- Location of Alignment in SEQ ID NO 353: from 1 to

40

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 354
- Ceres seq\_id 1976021
- Location of start within SEQ ID NO 352: at 239 nt.

45

- (C) Nomination and Annotation of Domains within

(Dp) Related Amino Acid Sequences

- Alignment No. 231
- gi No. 3688432

50

Predicted Polypeptide(s)

- Alignment No. 231
- gi No. 3688432

226

- Description: (AJ011705) 40S ribosomal protein S10 [Lumbricus rubellus]
- % Identity: 55.3
- Alignment Length: 164
- Location of Alignment in SEQ ID NO 354: from 1 to

5

115

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 355
- Ceres seq\_id 1976022
- Location of start within SEQ ID NO 352: at 287 nt.

10

- (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 232
- gi No. 3688432
- Description: (AJ011705) 40S ribosomal protein S10 [Lumbricus rubellus]
- % Identity: 55.3
- Alignment Length: 164
- Location of Alignment in SEQ ID NO 355: from 1 to

15

15

20

20

99

- (Ba) Polypeptide Activities: Extensin like protein activities and glycine rich protein activities.

25

## Maximum Length Sequence:

related to:

Clone IDs:

117263

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 356
- Ceres seq\_id 1976673
- Alternative transcription start site(s) located in SEQ ID NO 356:

-46,2,3,6,7

ID NO 356:

-46,2,3,6,7

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 357
- Ceres seq\_id 1976674
- Location of start within SEQ ID NO 356: at 3 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 357: at 34 aa.

40

45

- (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Ribosomal protein S21

- Location within SEQ ID NO 357: from 106 to 155

50

aa.



227

## (Dp) Related Amino Acid Sequences

- Alignment No. 233
- gi No. 1710750
- Description: 30S RIBOSOMAL PROTEIN S21 (BS-B)
- >gi11303814|dbj|BAA12470| (D84432) Yqex [Bacillus subtilis]
- >gi11890063|dbj|BAA12082| (D83717) Yqex [Bacillus subtilis]
- >gi12634987|emb|CAB14483| (Z99117) ribosomal protein S21 [Bacillus subtilis]

- % Identity: 38.5
- Alignment Length: 52
- Location of Alignment in SEQ ID NO 357: from 106 to 157

## (A) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 358
- Ceres seq\_id 1976675
- Location of start within SEQ ID NO 356: at 48 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 358: at 19 aa.

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- Ribosomal protein S21
- Location within SEQ ID NO 358: from 91 to 140 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 234
- gi No. 1710750
- Description: 30S RIBOSOMAL PROTEIN S21 (BS-B)
- >gi11303814|dbj|BAA12470| (D84432) Yqex [Bacillus subtilis]
- >gi11890063|dbj|BAA12082| (D83717) Yqex [Bacillus subtilis]
- >gi12634987|emb|CAB14483| (Z99117) ribosomal protein S21 [Bacillus subtilis]

- % Identity: 38.5
- Alignment Length: 52
- Location of Alignment in SEQ ID NO 358: from 91 to 142

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 359
- Ceres seq\_id 1976676
- Location of start within SEQ ID NO 356: at 192 nt.

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Ribosomal protein S21
- Location within SEQ ID NO 359: from 43 to 92 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 235
- gi No. 1710750

228

- Description: 30S RIBOSOMAL PROTEIN S21 (BS-B)
- >gi11303814|dbj|BAA12470| (D84432) Yqex [Bacillus subtilis]
- >gi11890063|dbj|BAA12082| (D83717) Yqex [Bacillus subtilis]
- >gi12634987|emb|CAB14483| (Z99117) ribosomal protein S21 [Bacillus subtilis]

- % Identity: 38.5
- Alignment Length: 52
- Location of Alignment in SEQ ID NO 359: from 43 to 94

## Maximum Length Sequence:

related to:  
Clone IDs:

42333

## Public Genomic DNA:

- gi No: 6041810
- Predicted Exons:
- INTR 93582 ... 94052
- INTR 94169 ... 94771

- INIT 93630 ... 94052
- TERM 94169 ... 94687
- gi No: 6091711
- Predicted Exons:
- INTR 60479 ... 60949
- INTR 61066 ... 61668

- INIT 60527 ... 60949
- TERM 61066 ... 61584
- gi No: 6102641
- Predicted Exons:
- INTR 52165 ... 52635
- INTR 52752 ... 53354

- INIT 52213 ... 52635
- TERM 52752 ... 53270
- gi No: 6453849
- Predicted Exons:
- INTR 59432 ... 58964
- INTR 58847 ... 58243

- INIT 59386 ... 58964
- TERM 58847 ... 58329
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 360
- Ceres seq\_id 2025186

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 361
- Ceres seq\_id 2025187

229  
- Location of start within SEQ ID NO 360: at 1 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- D-isomer specific 2-hydroxyacid dehydrogenases
- Location within SEQ ID NO 361: from 93 to 295 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 236
- gi No. 5458864
- Description: (AJ248287) GLYCERATE DEHYDROGENASE  
(Pyrococcus abyssi)
- % Identity: 42.8
- Alignment Length: 299
- Location of Alignment in SEQ ID NO 361: from 36

15 to 322

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 362
- Ceres seq\_id 2025188
- Location of start within SEQ ID NO 360: at 49 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- D-isomer specific 2-hydroxyacid dehydrogenases
- Location within SEQ ID NO 362: from 77 to 279 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 237
- gi No. 5458864
- Description: (AJ248287) GLYCERATE DEHYDROGENASE  
(Pyrococcus abyssi)
- % Identity: 42.8
- Alignment Length: 299
- Location of Alignment in SEQ ID NO 362: from 20

35 to 306

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 363
- Ceres seq\_id 2025189
- Location of start within SEQ ID NO 360: at 70 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- D-isomer specific 2-hydroxyacid dehydrogenases
- Location within SEQ ID NO 363: from 70 to 272 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 238
- gi No. 5458864

230

- Description: (AJ248287) GLYCERATE DEHYDROGENASE  
(Pyrococcus abyssi)

- % Identity: 42.8
- Alignment Length: 299
- Location of Alignment in SEQ ID NO 363: from 13

to 299

Maximum Length Sequence:

related to:

Clone IDs:

267004

Public Genomic DNA:

gi No: 2924730

Predicted Exons:

INTR 8331 ... 8882 OCKHMG-CDNA

SINGLE 0366 ... 8686 OCKHMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 364
- Ceres seq\_id 2025372
- Alternative transcription start site(s) located in SEQ ID NO 364: 21

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 365
- Ceres seq\_id 2025373
- Location of start within SEQ ID NO 364: at 1 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 239
- gi No. 132825
- Description: 50S RIBOSOMAL PROTEIN CL25,  
CHLOROPLAST PRECURSOR >gi171308|pit1|R5PM25 ribosomal protein  
pct125 precursor, chloroplast - garden pea  
>gi120877|emb1CBA321871 (X14022) psc125 ribosomal preprotein  
(AA -30 to 74) [Pisum sativum]
- % Identity: 56.1
- Alignment Length: 82
- Location of Alignment in SEQ ID NO 365: from 14

to 95

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 366
- Ceres seq\_id 2025374
- Location of start within SEQ ID NO 364: at 37 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 366: at 16 aa.

231

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 240

- gi No. 132925

- Description: 50S RIBOSOMAL PROTEIN CL25,

CHLOROPLAST PRECURSOR &gt;gi171308|pir1|IR5PM25 ribosomal protein

PscL25 precursor, chloroplast - garden pea

&gt;gi120877|emb|CAA321871 (X14022) PscL25 ribosomal prepro-pein

(AA -30 to 74) [Pisum sativum]

- % Identity: 56.1

- Alignment Length: 82

- Location of Alignment in SEQ ID NO 366: from 2 to

83

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 367

- Ceres seq\_id 2025375

- Location of start within SEQ ID NO 364: at 142 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 241

- gi No. 132825

- Description: 50S RIBOSOMAL PROTEIN CL25,

CHLOROPLAST PRECURSOR &gt;gi171308|pir1|IR5PM25 ribosomal protein

PscL25 precursor, chloroplast - garden pea

&gt;gi120877|emb|CAA321871 (X14022) PscL25 ribosomal prepro-pein

(AA -30 to 74) [Pisum sativum]

- % Identity: 56.1

- Alignment Length: 82

- Location of Alignment in SEQ ID NO 367: from 1 to

48

Maximum Length Sequence:

related to:

Clone IDs:

26930

Public Genomic DNA:

gi No: 5225383

Predicted Exons:

SINGLE 64539 ... 64276 GENBANK

INIT 64787 ... 64692 OCKHAMG-CDS

TERM 64620 ... 64276 OCKHAMG-CDS

gi No: 6136349

Predicted Exons:

INIT 126797... 126892 OCKHAMG-CDS

TERM 126964... 127308 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

232

- Pat. Appln. SEQ ID NO 368

- Ceres seq\_id 2025471

- Alternative transcription start site(s) located in SEQ

ID NO 368:

95,96,97,98,100,101,102,105,106,110

- Clone 26930 starts at 57 and ends at 631 in cDNA.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 369

- Ceres seq\_id 2025472

- Location of start within SEQ ID NO 368: at 1 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 242

- gi No. 5732082

- Description: (AF162444) contains similarity to

Solanum lycopersicum (tomato) wound induced protein

(GB:X59882) [Arabidopsis thaliana]

- % Identity: 100

- Alignment Length: 87

- Location of Alignment in SEQ ID NO 369: from 60

to 146

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 370

- Ceres seq\_id 2025473

- Location of start within SEQ ID NO 368: at 178 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 243

- gi No. 5732082

- Description: (AF162444) contains similarity to

Solanum lycopersicum (tomato) wound induced protein

(GB:X59882) [Arabidopsis thaliana]

- % Identity: 100

- Alignment Length: 87

- Location of Alignment in SEQ ID NO 370: from 1 to

87

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 371

- Ceres seq\_id 2025474

- Location of start within SEQ ID NO 368: at 214 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

233

- Alignment No. 244
- gi No. 5732082
- Description: (AF162444) contains similarity to Solanum lycopersicum (tomato) wound induced protein (GB:X59882) [Arabidopsis thaliana]
- % Identity: 100
- Alignment Length: 87
- Location of Alignment in SEQ ID NO 371: from 1 to

75

Maximum Length Sequence:  
related to:

C1one IDs:

27681

15 Public Genomic DNA:

gi No: 6434227

Predicted Exons:

INTR 51354 ... 51417 OCKHAMG-CDNA  
INTR 51529 ... 51786 OCKHAMG-CDNA

20

INTR 47952 ... 48020 OCKHAMG-CDS  
INTR 48144 ... 48440 OCKHAMG-CDS  
INTR 48542 ... 48757 OCKHAMG-CDS  
INTR 48908 ... 49027 OCKHAMG-CDS  
INTR 49117 ... 49654 OCKHAMG-CDS  
TERM 50075 ... 50184 OCKHAMG-CDS

25

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 372

- Ceres seq\_id 2025475

30 ID NO 372:  
- Alternative transcription start site(s) located in SEQ

1242,1251,1270,1271,1275,1276,1280,1287,1288,1290,1292,1293,1  
301,1302,1305,1307,1314,1318,1321,1325

35

- Clone 27681 starts at 1242 and ends at 2029 in cDNA.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 373

- Ceres seq\_id 2025476

- Location of start within SEQ ID NO 372: at 1 nt.

40

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 245

- gi No. 2739366

- Description: (AC002505) SFL6 like protein

[Arabidopsis thaliana]

- % Identity: 45

- Alignment Length: 373

50

234

to 418  
- Location of Alignment in SEQ ID NO 373: from 6-

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 374

- Ceres seq\_id 2025477

- Location of start within SEQ ID NO 372: at 316 nt.

5

(C) Nomination and Annotation of Domains within

10 Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 246

- gi No. 2739366

- Description: (AC002505) SFL6 like protein

[Arabidopsis thaliana]

- % Identity: 45

- Alignment Length: 373

- Location of Alignment in SEQ ID NO 374: from 1 to

15

313

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 375

- Ceres seq\_id 2025478

- Location of start within SEQ ID NO 372: at 364 nt.

- Location of Signal Peptide Cleavage Site within SEQ

25

ID NO 375: at 21 aa.

(C) Nomination and Annotation of Domains within

30 Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 247

- gi No. 2739366

- Description: (AC002505) SFL6 like protein

[Arabidopsis thaliana]

- % Identity: 45

- Alignment Length: 373

- Location of Alignment in SEQ ID NO 375: from 1 to

35

297

Maximum Length Sequence:

related to:

Clone IDs:

14564

Public Genomic DNA:

gi No: 4415905

Predicted Exons:

INTR 3235 ... 3181 OCKHAMG-CDNA

INTR 2694 ... 2619 OCKHAMG-CDNA

INTR 2509 ... 2408 OCKHAMG-CDNA

INTR 2319 ... 2171 OCKHAMG-CDNA

INTR 2076 ... 1784 OCKHAMG-CDNA

50

235

INIT 3191 ... 3191 OCKHAMG-CDS  
 INTR 2694 ... 2619 OCKHAMG-CDS  
 INTR 2509 ... 2408 OCKHAMG-CDS  
 INTR 2319 ... 2171 OCKHAMG-CDS  
 TERM 2076 ... 1920 OCKHAMG-CDS

gi No: 4581138

## Predicted Exons:

INTR 100911... 130857 OCKHAMG-CDNA  
 INTR 100370... 100295 OCKHAMG-CDNA  
 INTR 100185... 100084 OCKHAMG-CDNA  
 INTR 99995 ... 99847 OCKHAMG-CDNA  
 INTR 99752 ... 99460 OCKHAMG-CDNA

15

INIT 100867... 100857 OCKHAMG-CDS  
 INTR 100370... 100295 OCKHAMG-CDS  
 INTR 100185... 100084 OCKHAMG-CDS  
 INTR 99995 ... 99847 OCKHAMG-CDS  
 TERM 99752 ... 99596 OCKHAMG-CDS

20 (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 376
- Ceres seq\_id 2025524
- Alternative transcription start site(s) located in SEQ ID NO 376:

25

-35,-26,-25,-2,3,4,5,6,7,8,9,15,16,17,32

(3)

## Polypeptide Sequence

- Pat. Appln. SEQ ID NO 377
- Ceres seq\_id 2025525
- Location of start within SEQ ID NO 376: at 47 nt.

30

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein L24e
- Location within SEQ ID NO 377: from 3 to 73 aa.

35

## (Dp) Related Amino Acid Sequences

- Alignment No. 248
- gi No. 6094040
- Description: 60S RIBOSOMAL PROTEIN L24

>gi12961300|emb|CAA12358| (AJ225027) ribosomal protein L24  
 [Cicer arietinum]

- % Identity: 88.3

- Alignment Length: 163

- Location of Alignment in SEQ ID NO 377: from 1 to

45 163

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 378
- Ceres seq\_id 2025526
- Location of start within SEQ ID NO 376: at 212 nt.

50

236

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 249
- gi No. 6094040
- Description: 60S RIBOSOMAL PROTEIN L24

>gi12961300|emb|CAA12358| (AJ225027) ribosomal protein L24  
 [Cicer arietinum]

- % Identity: 88.3

- Alignment Length: 163

- Location of Alignment in SEQ ID NO 378: from 1 to

108

15 Maximum Length Sequence:

related to:

Clone IDs:

20767

Public Genomic DNA:

gi No: 4519195

## Predicted Exons:

INIT 50757 ... 50625 OCKHAMG-CDS  
 INTR 50001 ... 49847 OCKHAMG-CDS  
 INTR 49644 ... 49561 OCKHAMG-CDS  
 INTR 49391 ... 49236 OCKHAMG-CDS  
 INTR 49128 ... 49023 OCKHAMG-CDS  
 TERM 48810 ... 48662 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 379
- Ceres seq\_id 2025544
- Alternative transcription start site(s) located in SEQ ID NO 379:

656,672,680,685,701,703,706,709,713,717,728,736,741,743

- Clone 20767 starts at 656 and ends at 1623 in cDNA.

35

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 380
- Ceres seq\_id 2025545
- Location of start within SEQ ID NO 379: at 599 nt.

40

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Domain of unknown function
- Location within SEQ ID NO 380: from 134 to 247

45 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 250
- gi No. 1708292

50

237

- Description: HEAT-RESPONSIVE PROTEIN 12
- >gi11255116 (U50631) heat-responsive protein [Mus musculus]
- % Identity: 54.3
- Alignment Length: 129
- Location of Alignment in SEQ ID NO 380: from 119 to 247

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 381
- Ceres seq\_id 2025546
- Location of start within SEQ ID NO 379: at 1 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 381: at 50 aa.

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Zinc finger, C3HC4 type (RING finger)
- Location within SEQ ID NO 381: from 171 to 211

## (Dp) Related Amino Acid Sequences

- Alignment No. 251
- gi No. 2435518
- Description: (AF024504) contains similarity to C3HC4-type zinc fingers (Arabidopsis thaliana)
- % Identity: 31.4
- Alignment Length: 221
- Location of Alignment in SEQ ID NO 381: from 35 to 211

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 382
- Ceres seq\_id 2025547
- Location of start within SEQ ID NO 379: at 16 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 382: at 45 aa.

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Zinc finger, C3HC4 type (RING finger)
- Location within SEQ ID NO 382: from 166 to 206

## (Dp) Related Amino Acid Sequences

- Alignment No. 252
- gi No. 2435518
- Description: (AF024504) contains similarity to C3HC4-type zinc fingers (Arabidopsis thaliana)
- % Identity: 31.4
- Alignment Length: 221

238

- Location of Alignment in SEQ ID NO 382: from 30 to 206

## Maximum Length Sequence:

related to:  
Clone IDs:

7832

Public Genomic DNA:

gi No: 441592c

Predicted Exons:

INTR 60672 ... 61213 CCKHANG-CDNA

SINGLE 60712 ... 61119 CCKHANG-CDS

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 383
- Ceres seq\_id 2026207
- Alternative transcription start site(s) located in SEQ ID NO 383: 2,3,4,5

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 384
- Ceres seq\_id 2026208
- Location of start within SEQ ID NO 383: at 2 nt.

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 253
- gi No. 4757974
- Description: ref|NP\_004057.1|PCETN1 centrin, EF-hand protein, 1 >gi124934401sp|Q12798|CAT2\_HUMAN CALTRACTIN, ISOFORM 2 (CENTRIN) >gi1414993 (U03270) centrin [Homo sapiens]
- % Identity: 34.3
- Alignment Length: 102
- Location of Alignment in SEQ ID NO 384: from 39 to 120

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 385
- Ceres seq\_id 2026209
- Location of start within SEQ ID NO 383: at 41 nt.

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 254
- gi No. 4757974
- Description: ref|NP\_004057.1|PCETN1 centrin, EF-hand protein, 1 >gi124934401sp|Q12798|CAT2\_HUMAN CALTRACTIN,

239

ISOFORM 2 (CENTRIN) >gi1414993 (U03270) centrln [Homo sapiens]

- % Identity: 34.3
- Alignment Length: 102
- Location of Alignment in SEQ ID NO 385: from 6 to 107

5

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 386
- Ceres seq\_id 2026210
- Location of start within SEQ ID NO 383: at 62 nt.

10

(C) Nomination and Annotation of Domains within

(Dp) Related Amino Acid Sequences

- Alignment No. 255
- gi No. 4757974
- Description: refINP\_004057.1|pCETN1| centrln, EF-hand protein, 1 >gi12493440|sp|Q12798|CAT2\_HUMAN CALTRACTIN, ISOFORM 2 (CENTRIN) >gi1414993 (U03270) centrln [Homo sapiens]
- % Identity: 34.3
- Alignment Length: 102
- Location of Alignment in SEQ ID NO 386: from 1 to 107

25

Maximum Length Sequence:

Public Genomic DNA:

gi No: 4038029

Predicted Exons:

INIT	46188 ...	46125	GENBANK
TERM	46021 ...	45852	GENBANK
INTR	46258 ...	46125	OCKHANG-CDNA
INTR	46021 ...	45663	OCKHANG-CDNA
INIT	47157 ...	47119	OCKHANG-CDS
INTR	46278 ...	46125	OCKHANG-CDS
TERM	46021 ...	45852	OCKHANG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 387
- Ceres seq\_id 2026982

40

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 388
- Ceres seq\_id 2026983
- Location of start within SEQ ID NO 387: at 1 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 388: at 73 aa.

50

240

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Gamma-thionins family
- Location within SEQ ID NO 388: from 74 to 120 aa.

5

(Dp) Related Amino Acid Sequences

- Alignment No. 256
- gi No. 1209258
- Description: (L31937) protease inhibitor II [Brassica rapa]
- % Identity: 87
- Alignment Length: 77
- Location of Alignment in SEQ ID NO 388: from 44 to 120

10

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 389
- Ceres seq\_id 2326984
- Location of start within SEQ ID NO 387: at 34 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 389: at 54 aa.

20

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Gamma-thionins family
- Location within SEQ ID NO 389: from 63 to 109 aa.

25

(Dp) Related Amino Acid Sequences

- Alignment No. 257
- gi No. 1209258
- Description: (L31937) protease inhibitor II [Brassica rapa]
- % Identity: 87
- Alignment Length: 77
- Location of Alignment in SEQ ID NO 389: from 33 to 109

30

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 390
- Ceres seq\_id 2026985
- Location of start within SEQ ID NO 387: at 130 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 390: at 22 aa.

40

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Gamma-thionins family
- Location within SEQ ID NO 390: from 31 to 77 aa.

45

(Dp) Related Amino Acid Sequences

50

241

- Alignment No. 258
- gi No. 1209258
- Description: (L31937) protease inhibitor II (Brassica rapa)

- % Identity: 87
- Alignment Length: 77
- Location of Alignment in SEQ ID NO 390: from 1 to 77

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 391
- Ceres seq\_id 2026986
- Location of start within SEQ ID NO 387: at 30 nt.

## (C) Nomination and Annotation of Domains within

(Dp) Related Amino Acid Sequences

Maximum Length Sequence:

- related to:
- Clone IDs:

251127

Public Genomic DNA:

gi No: 4220643

Predicted Exons:

- |    |      |           |       |             |
|----|------|-----------|-------|-------------|
| 25 | INTR | 67833 ... | 68135 | OCKHANG-CDS |
|    | INTR | 68253 ... | 58366 | OCKHANG-CDS |
|    | INTR | 68457 ... | 68552 | OCKHANG-CDS |
| 30 | INTR | 68720 ... | 68794 | OCKHANG-CDS |
|    | INTR | 68898 ... | 68965 | OCKHANG-CDS |
|    | INTR | 69056 ... | 69066 | OCKHANG-CDS |
|    | INTR | 69422 ... | 69624 | OCKHANG-CDS |
|    | INTR | 69856 ... | 70202 | OCKHANG-CDS |
|    | INTR | 70315 ... | 70428 | OCKHANG-CDS |
| 35 | INTR | 70516 ... | 70611 | OCKHANG-CDS |
|    | INTR | 70923 ... | 70997 | OCKHANG-CDS |
|    | INTR | 71103 ... | 71173 | OCKHANG-CDS |
|    | TERM | 71737 ... | 71749 | OCKHANG-CDS |

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 392
- Ceres seq\_id 2027300
- Alternative transcription start site(s) located in SEQ ID NO 392: 38,53,55,63,82,84,85,103,160

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 393
- Ceres seq\_id 2027301
- Location of start within SEQ ID NO 392: at 177 nt.

50

242

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences

- Alignment No. 259
- gi No. 5020103
- Description: (AF148219) fibrillin (Nostoc PCC8009)

- % Identity: 30.3
- Alignment Length: 122
- Location of Alignment in SEQ ID NO 393: from 75 to 189

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 394
- Ceres seq\_id 2027302
- Location of start within SEQ ID NO 392: at 213 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- (Dp) Related Amino Acid Sequences

- Alignment No. 260
- gi No. 5020103
- Description: (AF148219) fibrillin (Nostoc PCC8009)

- % Identity: 30.3
- Alignment Length: 122
- Location of Alignment in SEQ ID NO 394: from 63 to 177

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 395
- Ceres seq\_id 2027303
- Location of start within SEQ ID NO 392: at 252 nt.

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- (Dp) Related Amino Acid Sequences
- Alignment No. 261
- gi No. 5020103
- Description: (AF148219) fibrillin (Nostoc PCC8009)

- % Identity: 30.3
- Alignment Length: 122
- Location of Alignment in SEQ ID NO 395: from 50 to 164

Maximum Length Sequence:

- related to:
- Clone IDs: 271260
- Public Genomic DNA:

50



- 243
- gi No: 6056186  
 Predicted Exons:  
 INIT 66600 ... 66465 OCKHANG-CDS  
 TERM 66390 ... 66083 OCKHANG-CDS
- 5 gi No: 6087919  
 Predicted Exons:  
 INIT 31657 ... 31522 OCKHANG-CDS  
 TERM 31447 ... 31140 OCKHANG-CDS
- 10 gi No: 619514  
 Predicted Exons:  
 INIT 32265 ... 32130 OCKHANG-CDS  
 TERM 32055 ... 31743 OCKHANG-CDS
- 15 gi No: 6143857  
 Predicted Exons:  
 INIT 34633 ... 34768 OCKHANG-CDS  
 TERM 34843 ... 35150 OCKHANG-CDS
- 20 gi No: 6175136  
 Predicted Exons:  
 INIT 34136 ... 34271 OCKHANG-CDS  
 TERM 34346 ... 34653 OCKHANG-CDS
- 25 gi No: 6223634  
 Predicted Exons:  
 INIT 34137 ... 34272 OCKHANG-CDS  
 TERM 34347 ... 34654 OCKHANG-CDS
- (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 396  
 - Ceres seq\_id 2027375
- 30 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 397  
 - Ceres seq\_id 2027376  
 - Location of start within SEQ ID NO 396: at 2 nt.
- 35 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 - Bacterial mutT protein  
 - Location within SEQ ID NO 397: from 27 to 68 aa.
- 40 (Dp) Related Amino Acid Sequences  
 - Alignment No. 262  
 - gi No. 2622420  
 - Description: (AE000895) mutator MutT protein  
 [Methanobacterium thermoautotrophicum]  
 - % Identity: 42.3  
 - Alignment Length: 98  
 - Location of Alignment in SEQ ID NO 397: from 13 to 109
- 45 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 398

- 244
- Ceres seq\_id 2027377  
 - Location of start within SEQ ID NO 396: at 17 nt.
- 5 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 - Bacterial mutT protein  
 - Location within SEQ ID NO 398: from 22 to 63 aa.
- 10 (Dp) Related Amino Acid Sequences  
 - Alignment No. 263  
 - gi No. 2622420  
 - Description: (AE000895) mutator MutT protein  
 [Methanobacterium thermoautotrophicum]  
 - % Identity: 42.3  
 - Alignment Length: 98  
 - Location of Alignment in SEQ ID NO 398: from 8 to 104
- 15 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 399  
 - Ceres seq\_id 2027378  
 - Location of start within SEQ ID NO 396: at 188 nt.
- 20 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 264  
 - gi No. 2622420  
 - Description: (AE000895) mutator MutT protein  
 [Methanobacterium thermoautotrophicum]  
 - % Identity: 42.3  
 - Alignment Length: 98  
 - Location of Alignment in SEQ ID NO 399: from 1 to 47
- 30 Maximum Length Sequence:  
 Public Genomic DNA:  
 gi No: 5868932  
 Predicted Exons:  
 INIT 99857 ... 96712 GENBANK  
 INTR 96123 ... 95883 GENBANK  
 TERM 95642 ... 95199 GENBANK
- 35 Maximum Length Sequence:  
 Public Genomic DNA:  
 gi No: 5868932  
 Predicted Exons:  
 INTR 96097 ... 95883 OCKHANG-CDNA  
 INTR 95642 ... 95047 OCKHANG-CDNA
- 40 INTR 96035 ... 95883 OCKHANG-CDS  
 TERM 95642 ... 95199 OCKHANG-CDS
- 45 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 400

- Ceres seq\_id 2028729 245

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 401
- Ceres seq\_id 2028730
- Location of start within SEQ ID NO 400: at 1 nt.

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- C2 domain
- Location within SEQ ID NO 401: from 671 to 762

aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 265
- gi No. 3142295
- Description: (AC002411) Strong similarity to phosphoribosylanthranilate transferase gb1086180 from Pisum sativum. [Arabidopsis thaliana]
- % Identity: 45.6
- Alignment Length: 865
- Location of Alignment in SEQ ID NO 401: from 211 to 1030

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 402
- Ceres seq\_id 2028731
- Location of start within SEQ ID NO 400: at 199 nt.

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- C2 domain
- Location within SEQ ID NO 402: from 605 to 696

35 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 266
- gi No. 3142295
- Description: (AC002411) Strong similarity to phosphoribosylanthranilate transferase gb1086180 from Pisum sativum. [Arabidopsis thaliana]
- % Identity: 45.6
- Alignment Length: 865
- Location of Alignment in SEQ ID NO 402: from 145 to 984

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 403
- Ceres seq\_id 2028732
- Location of start within SEQ ID NO 400: at 232 nt.

246

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- C2 domain
- Location within SEQ ID NO 403: from 594 to 685

5 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 267
- gi No. 3142295
- Description: (AC002411) Strong similarity to phosphoribosylanthranilate transferase gb1086180 from Pisum sativum. [Arabidopsis thaliana]
- % Identity: 45.6
- Alignment Length: 865
- Location of Alignment in SEQ ID NO 403: from 134 to 973

## Maximum Length Sequence:

related to:  
Clone IDs:

9214

- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 404
- Ceres seq\_id 2029079
- Alternative transcription start site(s) located in SEQ ID NO 404:  
-4,-1,2,3,4,8,9,10,11,14,15,16,17,18,20,23,24

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 405
- Ceres seq\_id 2029080
- Location of start within SEQ ID NO 404: at 3 nt.

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 268
- gi No. 132825
- Description: S05 RIBOSOMAL PROTEIN CL25, CHLOROPLAST PRECURSOR >gi171308|pir1|B5PM25 ribosomal protein >sc125 precursor, chloroplast - garden pea >gi120877|emo|CAA321871 (X14022) >sc125 ribosomal preprotein (AA -30 to 74) [Pisum sativum]
- % Identity: 56.1
- Alignment Length: 82
- Location of Alignment in SEQ ID NO 405: from 14 to 95

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 406

247

- Ceres seq\_id 2029081
- Location of start within SEQ ID NO 404: at 39 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 406: at 16 aa.

5

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 269
- gi No. 132825
- Description: 50S RIBOSOMAL PROTEIN CL25, CHLOROPLAST PRECURSOR >gi|71308|pir|IRSPM25 ribosomal protein PsCL25 precursor, chloroplast - garden pea
- >gi|20877|emb|CAA321871 (X14022) PsCL25 ribosomal preprotein (AA -30 to 74) [Pisum sativum]
- % Identity: 56.1
- Alignment Length: 82
- Location of Alignment in SEQ ID NO 406: from 2 to

20

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 407
- Ceres seq\_id 2029032
- Location of start within SEQ ID NO 404: at 144 nt.

25

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 273
- gi No. 132825
- Description: 50S RIBOSOMAL PROTEIN CL25, CHLOROPLAST PRECURSOR >gi|71308|pir|IRSPM25 ribosomal protein PsCL25 precursor, chloroplast - garden pea
- >gi|20877|emb|CAA321871 (X14022) PsCL25 ribosomal preprotein (AA -30 to 74) [Pisum sativum]
- % Identity: 56.1
- Alignment Length: 82
- Location of Alignment in SEQ ID NO 407: from 1 to

48

## Maximum Length Sequence:

## Public Genomic DNA:

gi No: 6434227

## Predicted Exons:

- INTR 50076 ... 50370 OCKHAMG-CDNA
- INTR 50793 ... 50874 OCKHAMG-CDNA
- INTR 51163 ... 51227 OCKHAMG-CDNA
- INTR 51331 ... 51417 OCKHAMG-CDNA
- INTR 51529 ... 51786 OCKHAMG-CDNA
- INIT 50200 ... 50370 OCKHAMG-CDS

50

248

- INTR 50793 ... 50874 OCKHAMG-CDS
- INTR 51163 ... 51227 OCKHAMG-CDS
- INTR 51331 ... 51417 OCKHAMG-CDS
- TERM 51529 ... 51630 OCKHAMG-CDS

5

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 408
- Ceres seq\_id 2029225

10

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 409
- Ceres seq\_id 2029226
- Location of start within SEQ ID NO 408: at 125 nt.

15

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 271
- gi No. 1352051
- Description: ATP SYNTHASE D CHAIN, MITOCHONDRIAL >gi|484646|pir|J50739 H+-transporting ATP synthase (EC 3.6.1.34) chain d precursor - rat >gi|286204|dbj|BAA02422| (D13120) ATP synthase subunit d precursor [Rattus norvegicus]
- % Identity: 28.6
- Alignment Length: 145
- Location of Alignment in SEQ ID NO 409: from 15 to 154

20

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 410
- Ceres seq\_id 2029227
- Location of start within SEQ ID NO 408: at 194 nt.

30

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 272
- gi No. 1352051
- Description: ATP SYNTHASE D CHAIN, MITOCHONDRIAL >gi|484646|pir|J50739 H+-transporting ATP synthase (EC 3.6.1.34) chain d precursor - rat >gi|286204|dbj|BAA02422| (D13120) ATP synthase subunit d precursor [Rattus norvegicus]
- % Identity: 28.6
- Alignment Length: 145
- Location of Alignment in SEQ ID NO 410: from 1 to 131

40

## Maximum Length Sequence:

## Public Genomic DNA:

gi No: 6056185

50

249

## Predicted Exons:

INTR 2155 ... 22692 OCKHAMG-CDNA  
INTR 22891 ... 23339 OCKHAMG-CDNA

INTR 22502 ... 22692 OCKHAMG-CDNA  
INTR 22891 ... 23338 OCKHAMG-CDNA

INTR 21706 ... 22692 OCKHAMG-CDS  
TERM 22891 ... 23124 OCKHAMG-CDS

gi No: 6091703

## Predicted Exons:

INTR 21555 ... 22692 OCKHAMG-CDNA  
INTR 22891 ... 23339 OCKHAMG-CDNA

INTR 22502 ... 22692 OCKHAMG-CDNA  
INTR 22891 ... 23338 OCKHAMG-CDNA

INTR 21706 ... 22692 OCKHAMG-CDS  
TERM 22891 ... 23124 OCKHAMG-CDS

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 411  
- Ceres seq\_id 2029422

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 412  
- Ceres seq\_id 2029423

- Location of start within SEQ ID NO 411: at 152 nt.

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- CONSTANS family zinc finger  
- Location within SEQ ID NO 412: from 17 to 57 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 273  
- gi No. 1076301

40 thaliana >gi1161514|emb|CAA644071 (X94937) CONSTANS protein [Arabidopsis thaliana]

- % Identity: 49.2  
- Alignment Length: 63  
- Location of Alignment in SEQ ID NO 412: from 12

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 413  
- Ceres seq\_id 2029424

50 - Location of start within SEQ ID NO 411: at 155 nt.

250

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- CONSTANS family zinc finger  
- Location within SEQ ID NO 413: from 16 to 56 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 274  
- gi No. 1076301

10 thaliana >gi1161514|emb|CAA644071 (X94937) CONSTANS protein [Arabidopsis thaliana]

- % Identity: 49.2  
- Alignment Length: 63  
- Location of Alignment in SEQ ID NO 413: from 11

## (B) Polypeptide Sequence

15 to 73  
- Pat. Appln. SEQ ID NO 414  
- Ceres seq\_id 2029425  
- Location of start within SEQ ID NO 411: at 443 nt.

## (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences  
- Alignment No. 275  
- gi No. 1076301

25 thaliana >gi1161514|emb|CAA644071 (X94937) CONSTANS protein [Arabidopsis thaliana]

30 - % Identity: 69.8  
- Alignment Length: 43  
- Location of Alignment in SEQ ID NO 414: from 260 to 302

## (Dp) Maximum Length Sequence: Public Genomic DNA:

gi No: 5832738

35  
40  
INTR 56241 ... 56150 OCKHAMG-CDNA  
INTR 56041 ... 55864 OCKHAMG-CDNA  
INTR 55493 ... 55317 OCKHAMG-CDNA  
INTR 55216 ... 55083 OCKHAMG-CDNA  
INTR 54899 ... 54803 OCKHAMG-CDNA  
INTR 54725 ... 54609 OCKHAMG-CDNA  
INTR 54510 ... 54415 OCKHAMG-CDNA  
INTR 54288 ... 54176 OCKHAMG-CDNA  
INTR 54077 ... 53779 OCKHAMG-CDNA

45  
50  
INTR 56181 ... 56150 OCKHAMG-CDS  
INTR 56041 ... 55864 OCKHAMG-CDS  
INTR 55493 ... 55317 OCKHAMG-CDS

251  
INTR 55216 ... 55085 OCKHAMG-CDS  
INTR 54898 ... 54803 OCKHAMG-CDS  
INTR 54725 ... 54609 OCKHAMG-CDS  
INTR 54510 ... 54415 OCKHAMG-CDS  
INTR 54288 ... 54176 OCKHAMG-CDS  
TERM 54077 ... 53954 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 415  
- Ceres seq\_id 2029806

10

(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 416  
- Ceres seq\_id 2029807  
- Location of start within SEQ ID NO 415: at 1 nt.

15

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
- Aminotransferase class IV  
- Location within SEQ ID NO 416: from 68 to 352 aa.

20

(Dp) Related Amino Acid Sequences  
- Alignment No. 276  
- gi No. 3540183  
- Description: (AC004122) Highly Similar to  
branched-chain amino acid aminotransferase [Arabidopsis  
thaliana]  
- % Identity: 53.8  
- Alignment Length: 318  
- Location of Alignment in SEQ ID NO 416: from 55  
to 371

25

(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 417  
- Ceres seq\_id 2029808  
- Location of start within SEQ ID NO 415: at 61 nt.

30

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
- Aminotransferase class IV  
- Location within SEQ ID NO 417: from 48 to 332 aa.

35

(Dp) Related Amino Acid Sequences  
- Alignment No. 277  
- gi No. 3540183  
- Description: (AC004122) Highly Similar to  
branched-chain amino acid aminotransferase [Arabidopsis  
thaliana]  
- % Identity: 53.8  
- Alignment Length: 318

40

45

50

252  
- Location of Alignment in SEQ ID NO 417: from 35  
to 351

5

(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 418  
- Ceres seq\_id 2029809  
- Location of start within SEQ ID NO 415: at 163 nt.

10

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
- Aminotransferase class IV  
- Location within SEQ ID NO 418: from 14 to 298 aa.

15

(Dp) Related Amino Acid Sequences  
- Alignment No. 278  
- gi No. 3540183  
- Description: (AC004122) Highly Similar to  
branched-chain amino acid aminotransferase [Arabidopsis  
thaliana]  
- % Identity: 53.8  
- Alignment Length: 312  
- Location of Alignment in SEQ ID NO 418: from 1 to  
317

20

25

Maximum Length Sequence:  
Public Genomic DNA:  
gi No: 5822667  
Predicted Exons:  
INTR 59013 ... 59999 OCKHAMG-CDNA  
SINGLE 59071 ... 59769 OCKHAMG-CDS  
gi No: 6041831  
Predicted Exons:  
INTR 53943 ... 54929 OCKHAMG-CDNA  
SINGLE 54001 ... 54599 OCKHAMG-CDS

30

35

(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 419  
- Ceres seq\_id 2030038

40

(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 420  
- Ceres seq\_id 2030039  
- Location of start within SEQ ID NO 419: at 2 nt.

45

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 279

50

253

- gi No. 4009495  
- Description: (AF068904) YlmG [Staphylococcus aureus]

5  
to 239  
- % Identity: 31.7  
- Alignment Length: 63  
- Location of Alignment in SEQ ID NO 420: from 177

## (B) Polypeptide Sequence

10  
- Pat. Appln. SEQ ID NO 421  
- Ceres seq\_id 2C30040  
- Location of start within SEQ ID NO 419: at 59 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 421: at 19 aa.

## (C) Nomination and Annotation of Domains within

15  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences

20  
- Alignment No. 280  
- gi No. 4009495  
- Description: (AF068904) YlmG [Staphylococcus aureus]

25  
- % Identity: 31.7  
- Alignment Length: 63  
- Location of Alignment in SEQ ID NO 421: from 158 to 220

## Maximum Length Sequence:

Public Genomic DNA:

gi No: 4539290

## Predicted Exons:

30  
INIT 6134 ... 6132 GENBANK  
INTR 6021 ... 5973 GENBANK  
INTR 5588 ... 5309 GENBANK  
INTR 5226 ... 5106 GENBANK  
TERM 4951 ... 4871 GENBANK

35  
INTR 6208 ... 6132 OCKHAMG-CDNA  
INTR 6021 ... 5973 OCKHAMG-CDNA  
INTR 5588 ... 5309 OCKHAMG-CDNA  
INTR 5226 ... 5106 OCKHAMG-CDNA  
INTR 4951 ... 4724 OCKHAMG-CDNA

40  
INTR 6134 ... 6132 OCKHAMG-CDS  
INTR 6021 ... 5973 OCKHAMG-CDS  
INTR 5588 ... 5309 OCKHAMG-CDS  
INTR 5226 ... 5106 OCKHAMG-CDS  
TERM 4951 ... 4871 OCKHAMG-CDS

45  
gi No: 4914454  
Predicted Exons:  
INTR 39557 ... 39481 OCKHAMG-CDNA

254

INTR 39370 ... 39322 OCKHAMG-CDNA  
INTR 38937 ... 38658 OCKHAMG-CDNA  
INTR 38575 ... 38455 OCKHAMG-CDNA  
INTR 38300 ... 38073 OCKHAMG-CDNA

5  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 422  
- Ceres seq\_id 2031778

## (B) Polypeptide Sequence

10  
- Pat. Appln. SEQ ID NO 423  
- Ceres seq\_id 2031779  
- Location of start within SEQ ID NO 422: at 3 nt.

## (C) Nomination and Annotation of Domains within

15  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences

20  
- Alignment No. 281  
- gi No. 3688432  
- Description: (AJ011705) 40S ribosomal protein S10 [Lumbricus rubellus]  
- % Identity: 53.1

25  
- Alignment Length: 164  
- Location of Alignment in SEQ ID NO 423: from 25 to 186

## (B) Polypeptide Sequence

30  
- Pat. Appln. SEQ ID NO 424  
- Ceres seq\_id 2C31780  
- Location of start within SEQ ID NO 422: at 75 nt.

35  
(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences

- Alignment No. 282  
- gi No. 3688432  
- Description: (AJ011705) 40S ribosomal protein S10 [Lumbricus rubellus]  
- % Identity: 53.1

40  
- Alignment Length: 164  
- Location of Alignment in SEQ ID NO 424: from 1 to 162

## (B) Polypeptide Sequence

45  
- Pat. Appln. SEQ ID NO 425  
- Ceres seq\_id 2031781  
- Location of start within SEQ ID NO 422: at 210 nt.

50  
(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

255  
(Dp) Related Amino Acid Sequences  
- Alignment No. 283  
- gi No. 3688432  
- Description: (AJ011705) 40S ribosomal protein S10  
[Lumbricus rubellus]  
- % Identity: 53.1  
- Alignment Length: 164  
- Location of Alignment in SEQ ID NO 425: from 1 to 117

10  
Maximum Length Sequence:  
Public Genomic DNA:  
gi No: 6434226  
Predicted Exons:  
INTR 17873 ... 17514 OCKHAMG-CDNA  
INTR 17046 ... 16869 OCKHAMG-CDNA  
INIT 17854 ... 17514 OCKHAMG-CDS  
TERM 17046 ... 17028 OCKHAMG-CDS

20  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 426  
- Ceres seq\_id 2032723

25  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 427  
- Ceres seq\_id 2032724  
- Location of start within SEQ ID NO 426: at 20 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 427: at 24 aa.

30  
(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
- Plant lipid transfer protein family  
- Location within SEQ ID NO 427: from 28 to 115 aa.

35  
(Dp) Related Amino Acid Sequences  
- Alignment No. 284  
- gi No. 899224  
- Description: (X60318) E2 [Brassica napus]  
>gi|384340|prf||1905428A phospholipid transfer protein [Brassica napus]  
- % Identity: 79  
- Alignment Length: 119  
- Location of Alignment in SEQ ID NO 427: from 1 to 119

40  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 428  
- Ceres seq\_id 2032725

256  
- Location of start within SEQ ID NO 426: at 71 nt.

5  
(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
- Plant lipid transfer protein family  
- Location within SEQ ID NO 428: from 11 to 98 aa.

10  
(Dp) Related Amino Acid Sequences  
- Alignment No. 285  
- gi No. 899224  
- Description: (X60318) E2 [Brassica napus]  
>gi|384340|prf||1905428A phospholipid transfer protein [Brassica napus]  
- % Identity: 79  
- Alignment Length: 119  
- Location of Alignment in SEQ ID NO 428: from 1 to 102

15  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 429  
- Ceres seq\_id 2032726  
- Location of start within SEQ ID NO 426: at 1 nt.

20  
(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
Maximum Length Sequence:  
Public Genomic DNA:  
gi No: 6143856  
Predicted Exons:  
INIT 30571 ... 30464 OCKHAMG-CDS  
TERM 30327 ... 30238 OCKHAMG-CDS  
gi No: 6223633  
Predicted Exons:  
INIT 30571 ... 30464 OCKHAMG-CDS  
TERM 30327 ... 30238 OCKHAMG-CDS

35  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 430  
- Ceres seq\_id 2035536

40  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 431  
- Ceres seq\_id 2035537  
- Location of start within SEQ ID NO 430: at 1 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 431: at 20 aa.

45  
ID NO 431: at 20 aa.

50

257

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 286

- gi No. 3024192

SUBUNIT (COMPLEX I-MFE) (CI-MFE) >gi12275037|emb|CAA689771  
(Y07708) NADH dehydrogenase [Mus musculus]

- % Identity: 31.5

- Alignment Length: 54

- Location of Alignment in SEQ ID NO 431: from 4 to

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 432

- Ceres seq\_id 2035538

- Location of start within SEQ ID NO 430: at 2 nt.

(C) Nomination and Annotation of Domains within

(Dp) Related Amino Acid Sequences

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 433

- Ceres seq\_id 2035538

- Location of start within SEQ ID NO 430: at 25 nt.

ID NO 433: at 17 aa.

(C) Nomination and Annotation of Domains within

(Dp) Related Amino Acid Sequences

- Alignment No. 287

- gi No. 3024192

SUBUNIT (COMPLEX I-MFE) (CI-MFE) >gi12275037|emb|CAA689771  
(Y07708) NADH dehydrogenase [Mus musculus]

- % Identity: 31.5

- Alignment Length: 54

- Location of Alignment in SEQ ID NO 433: from 1 to

Maximum Length Sequence:

Public Genomic DNA:

gi No: 6136357

Predicted Exons:

INIT	51922 ...	51830	OCKHAMG-CDS
INTR	51129 ...	51055	OCKHAMG-CDS
INTR	50958 ...	50806	OCKHAMG-CDS
INTR	50706 ...	50608	OCKHAMG-CDS
TERM	50326 ...	50243	OCKHAMG-CDS

258

gi No: 6143858

Predicted Exons:

INIT 51922 ...

INTR 51129 ...

INTR 50958 ...

INTR 50706 ...

TERM 50326 ...

gi No: 6149510

Predicted Exons:

INTR 9035 ...

INTR 9208 ...

INTR 9460 ...

INTR 9840 ...

(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 434  
- Ceres seq\_id 2035575

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 435

- Ceres seq\_id 2035576

- Location of start within SEQ ID NO 434: at 1 nt.

(C) Nomination and Annotation of Domains within

(Dp) Related Amino Acid Sequences

- Alignment No. 288

- gi No. 6066609

- Description: (A000987) chloroplast channel  
forming outer membrane protein [Pisum sativum]

- % Identity: 54.2

- Alignment Length: 177

- Location of Alignment in SEQ ID NO 435: from 1 to

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 436

- Ceres seq\_id 2035577

- Location of start within SEQ ID NO 434: at 13 nt.

(C) Nomination and Annotation of Domains within

(Dp) Related Amino Acid Sequences

- Alignment No. 289

- gi No. 6066609

- Description: (A000987) chloroplast channel  
forming outer membrane protein [Pisum sativum]

- % Identity: 54.2

- Alignment Length: 177



259  
- Location of Alignment in SEQ ID NO 436: from 1 to 163

## (B) Polypeptide Sequence

- 5 - Pat. Appln. SEQ ID NO 437  
- Ceres seq\_id 2035578  
- Location of start within SEQ ID NO 434: at 40 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 10 - Alignment No. 290  
- gi No. 6066609  
- Description: (AJ009987) chloroplast channel forming outer membrane protein (Pisum sativum)  
15 - % Identity: 54.2  
- Alignment Length: 177  
- Location of Alignment in SEQ ID NO 437: from 1 to 154

## Maximum Length Sequence:

## Public Genomic DNA:

gi No: 6434213

## Predicted Exons:

25 INIT 54056 ... 54085 OCKHAMG-CDS  
INTR 54178 ... 54325 OCKHAMG-CDS  
INTR 54401 ... 54572 OCKHAMG-CDS  
INTR 54807 ... 54993 OCKHAMG-CDS  
TERM 55079 ... 55291 OCKHAMG-CDS

## (Ac) cDNA Polynucleotide Sequence

- 30 - Pat. Appln. SEQ ID NO 438  
- Ceres seq\_id 2036457

## (B) Polypeptide Sequence

- 35 - Pat. Appln. SEQ ID NO 439  
- Ceres seq\_id 2036458  
- Location of start within SEQ ID NO 438: at 1 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- 40 - Ribosomal protein S3, C-terminal domain.  
- Location within SEQ ID NO 439: from 104 to 188

45 aa.

## (Dp) Related Amino Acid Sequences

- 45 - Alignment No. 291  
- gi No. 133940  
- Description: 40S RIBOSOMAL PROTEIN S3A (SLA)  
50 >gi|70851|pir|IR3XL3A ribosomal protein S3a - African clawed

260  
frog >gi|65091|emb|CAA40592| (X57322) ribosomal protein Sla [Xenopus laevis]

- 5 - % Identity: 75.1  
- Alignment Length: 250  
- Location of Alignment in SEQ ID NO 439: from 1 to 249

## (B) Polypeptide Sequence

- 10 - Pat. Appln. SEQ ID NO 440  
- Ceres seq\_id 2036459  
- Location of start within SEQ ID NO 438: at 130 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- 15 - Ribosomal protein S3, C-terminal domain.  
- Location within SEQ ID NO 440: from 61 to 145 aa.

## (Dp) Related Amino Acid Sequences

- 20 - Alignment No. 292  
- gi No. 133940  
- Description: 40S RIBOSOMAL PROTEIN S3A (SLA)  
>gi|70851|pir|IR3XL3A ribosomal protein S3a - African clawed frog >gi|65091|emb|CAA40592| (X57322) ribosomal protein Sla [Xenopus laevis]  
25 - % Identity: 75.1  
- Alignment Length: 250  
- Location of Alignment in SEQ ID NO 440: from 1 to 206

## Maximum Length Sequence:

## Public Genomic DNA:

gi No: 6434218

## Predicted Exons:

35 INIT 85292 ... 85292 OCKHAMG-CDS  
INTR 84903 ... 84828 OCKHAMG-CDS  
INTR 84730 ... 84629 OCKHAMG-CDS  
INTR 84524 ... 84376 OCKHAMG-CDS  
INTR 84282 ... 84156 OCKHAMG-CDS  
TERM 84148 ... 84065 OCKHAMG-CDS

## (Ac) cDNA Polynucleotide Sequence

- 40 - Pat. Appln. SEQ ID NO 441  
- Ceres seq\_id 2036585

## (B) Polypeptide Sequence

- 45 - Pat. Appln. SEQ ID NO 442  
- Ceres seq\_id 2036586  
- Location of start within SEQ ID NO 441: at 1 nt.

261

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- Ribosomal protein L24e
- Location within SEQ ID NO 442: from 3 to 73 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 293
- gi No. 6094040
- Description: 60S RIBOSOMAL PROTEIN L24
- >gi|29613001emb|CAA123581 (AJ225027) ribosomal protein L24
- [Cicer arietinum]
- % Identity: 86.6
- Alignment Length: 158
- Location of Alignment in SEQ ID NO 442: from 1 to

- 15 156
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 443
  - Ceres seq\_id 2036587
  - Location of start within SEQ ID NO 441: at 166 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 294
- gi No. 6094040
- Description: 60S RIBOSOMAL PROTEIN L24
- >gi|29613001emb|CAA123581 (AJ225027) ribosomal protein L24
- [Cicer arietinum]
- % Identity: 86.6
- Alignment Length: 158
- Location of Alignment in SEQ ID NO 443: from 1 to

- 30 101
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 443
  - Ceres seq\_id 2036587
  - Location of start within SEQ ID NO 441: at 166 nt.

35 Maximum Length Sequence:  
Public Genomic DNA:

gi No: 3668776

- 40
- INIT 19552 ... 19704 OCKHANG-CDS
- TERM 19947 ... 20279 OCKHANG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 444
- Ceres seq\_id 2039554

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 445
- Ceres seq\_id 2039555
- Location of start within SEQ ID NO 444: at 1 nt.

262

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- Ribosomal protein L24e
- Location within SEQ ID NO 442: from 3 to 73 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 295
- gi No. 5882720
- Description: (AC008263) Similar to gb|D86180
- phosphoribosylanthranilate transferase from Pisum sativum and
- contains 2 PF00168 C2 (phospholipid binding) domains. ESTs
- gb|H76726, gb|I45544 and gb|N96377 come from this gene.
- [Arab...]
- % Identity: 99.3
- Alignment Length: 151
- Location of Alignment in SEQ ID NO 445: from 1 to

- 15 151
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 446
  - Ceres seq\_id 2039556
  - Location of start within SEQ ID NO 444: at 136 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 296
- gi No. 5882720
- Description: (AC008263) Similar to gb|D86180
- phosphoribosylanthranilate transferase from Pisum sativum and
- contains 2 PF00168 C2 (phospholipid binding) domains. ESTs
- gb|H76726, gb|I45544 and gb|N96377 come from this gene.
- [Arab...]
- % Identity: 99.3
- Alignment Length: 151
- Location of Alignment in SEQ ID NO 446: from 1 to

- 30 106
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 447
  - Ceres seq\_id 2044283
  - Location of start within SEQ ID NO 444: at 1 nt.

35 Maximum Length Sequence:  
Public Genomic DNA:

gi No: 4263586

- 40
- INIT 44660 ... 44729 OCKHANG-CDS
- INTR 44822 ... 44985 OCKHANG-CDS
- INTR 45066 ... 45099 OCKHANG-CDS
- INTR 45171 ... 45264 OCKHANG-CDS
- TERM 45352 ... 45367 OCKHANG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 447
- Ceres seq\_id 2044283

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 447
- Ceres seq\_id 2044283
- Location of start within SEQ ID NO 444: at 1 nt.

263

- 5 ID NO 448: at 19 aa.
- Pat. Appln. SEQ ID NO 448
  - Ceres seq\_id 2044284
  - Location of start within SEQ ID NO 447: at 1 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 448: at 19 aa.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 10
- Alignment No. 297
  - gi No. 4512613
  - Description: (AC004793) F28K20.12 [Arabidopsis thaliana]

- % Identity: 99.2

- 15
- Alignment Length: 122
  - Location of Alignment in SEQ ID NO 448: from 1 to 122

## (B) Polypeptide Sequence

- 20
- Pat. Appln. SEQ ID NO 449

- Ceres seq\_id 2044285
- Location of start within SEQ ID NO 447: at 13 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 25
- Alignment No. 298
  - gi No. 4512613
  - Description: (AC004793) F28K20.12 [Arabidopsis thaliana]

- % Identity: 99.2

- 30
- Alignment Length: 122
  - Location of Alignment in SEQ ID NO 448: from 1 to 118

## (B) Polypeptide Sequence

- 35
- Pat. Appln. SEQ ID NO 450

- Ceres seq\_id 2044286
- Location of start within SEQ ID NO 447: at 16 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 40
- Alignment No. 299
  - gi No. 4512613
  - Description: (AC004793) F28K20.12 [Arabidopsis thaliana]

- % Identity: 99.2

- 45
- Alignment Length: 122
  - Location of Alignment in SEQ ID NO 450: from 1 to 117

264

## Maximum Length Sequence:

related to:

Clone IDs:

- 5
- 157859

Public Genomic DNA:

gi No: 4519187

Predicted Exons:

- 10
- INIT 64212 ... 64316 OCKHAMG-CDS
  - INTR 64435 ... 64599 OCKHAMG-CDS
  - INTR 64655 ... 64789 OCKHAMG-CDS
  - INTR 64869 ... 64952 OCKHAMG-CDS
  - INTR 65028 ... 65143 OCKHAMG-CDS
  - INTR 65245 ... 65346 OCKHAMG-CDS
  - INTR 65417 ... 65530 OCKHAMG-CDS
  - INTR 65621 ... 65677 OCKHAMG-CDS
  - INTR 65790 ... 65856 OCKHAMG-CDS
  - INTR 65929 ... 65982 OCKHAMG-CDS
  - INTR 66089 ... 66163 OCKHAMG-CDS
  - INTR 66558 ... 66593 OCKHAMG-CDS
  - INTR 66699 ... 66811 OCKHAMG-CDS
  - TERM 68086 ... 68098 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- 25
- Pat. Appln. SEQ ID NO 451
  - Ceres seq\_id 2048114
  - Alternative transcription start site(s) located in SEQ ID NO 451:
  - 456
    - Clone 157859 starts at 456 and ends at in cDNA.

## (B) Polypeptide Sequence

- 30
- Pat. Appln. SEQ ID NO 452
  - Ceres seq\_id 2048115
  - Location of start within SEQ ID NO 451: at 1 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 452: at 34 aa.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 40
- Alignment No. 300
  - gi No. 4830384
  - Description: (AC007357) F3F19.3 [Arabidopsis thaliana]

- % Identity: 66.2

- 45
- Alignment Length: 201
  - Location of Alignment in SEQ ID NO 452: from 95 to 295

## (B) Polypeptide Sequence

- 50
- Pat. Appln. SEQ ID NO 453

265

- Ceres seq\_id 2048116
- Location of start within SEQ ID NO 451: at 19 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 453: at 28 aa.

5 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 301
- gi No. 4850384
- Description: (AC007357) EF19.3 [Arabidopsis thaliana]

- % Identity: 66.2

- Alignment Length: 201

- Location of Alignment in SEQ ID NO 453: from 89 to 289

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 454
- Ceres seq\_id 2048117
- Location of start within SEQ ID NO 451: at 22 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 454: at 27 aa.

25 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 302
- gi No. 4850384
- Description: (AC007357) EF19.3 [Arabidopsis thaliana]

- % Identity: 66.2

- Alignment Length: 201

- Location of Alignment in SEQ ID NO 454: from 88 to 288

Maximum Length Sequence:

Public Genomic DNA:

gi No: 4319194

Predicted Exons:

- |      |           |       |              |
|------|-----------|-------|--------------|
| INTR | 15272 ... | 15482 | OCKHANG-CDNA |
| INTR | 15570 ... | 15615 | OCKHANG-CDNA |
| INTR | 16443 ... | 16834 | OCKHANG-CDNA |
| INTR | 15350 ... | 15482 | OCKHANG-CDS  |
| INTR | 15570 ... | 15615 | OCKHANG-CDS  |
| TERM | 16443 ... | 16671 | OCKHANG-CDS  |

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 455
- Ceres seq\_id 2048271

266

- 5 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 456
  - Ceres seq\_id 2048272
  - Location of start within SEQ ID NO 455: at 79 nt.

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Ribosomal protein S16
- Location within SEQ ID NO 456: from 9 to 70 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 303
- gi No. 3861401
- Description: (AJ235273) 30S RIBOSOMAL PROTEIN S16 (rpSP) [Rickettsia prowazekii]

- % Identity: 43.7

- Alignment Length: 103
- Location of Alignment in SEQ ID NO 456: from 1 to 102

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 457
- Ceres seq\_id 2048273
- Location of start within SEQ ID NO 455: at 139 nt.

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Ribosomal protein S16
- Location within SEQ ID NO 457: from 1 to 50 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 304
- gi No. 3861401
- Description: (AJ235273) 30S RIBOSOMAL PROTEIN S16 (rpSP) [Rickettsia prowazekii]

- % Identity: 43.7

- Alignment Length: 103
- Location of Alignment in SEQ ID NO 457: from 1 to 32

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 458
- Ceres seq\_id 2048274
- Location of start within SEQ ID NO 455: at 232 nt.

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Alignment No. 305

267

- 5
- gi No. 3861401
  - Description: (AJ235273) 30S RIBOSOMAL PROTEIN S16 (rpsP) [Rickettsia prowazekii]
  - % Identity: 43.7
  - Alignment Length: 103
  - Location of Alignment in SEQ ID NO 458: from 1 to

51

## Maximum Length Sequence:

Public Genomic DNA:

gi No: 4519195

## Predicted Exons:

INIT	48116 ...	48096	OCKHANG-CDS
INTR	48033 ...	47989	OCKHANG-CDS
INTR	47604 ...	47531	OCKHANG-CDS
INTR	47435 ...	47386	OCKHANG-CDS
TERM	47286 ...	47192	OCKHANG-CDS

## (Ac) cDNA Polynucleotide Sequence

- 20
- Pat. Appln. SEQ ID NO 459
  - Ceres seq\_id 2048331

## (B) Polypeptide Sequence

- 25
- Pat. Appln. SEQ ID NO 460
  - Ceres seq\_id 2048332
  - Location of start within SEQ ID NO 459: at 1 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- 30
- Domain of unknown function
  - Location within SEQ ID NO 460: from 8 to 92 aa.

## (Dp) Related Amino Acid Sequences

- 35
- Alignment No. 306
  - gi No. 5032215
  - Description: ref|NP\_005827.1|pUK114| translational inhibitor protein p14.5
  - >gi|1717975|sp|P52758|UK14\_HUMAN 14.5 KD TRANSLATIONAL INHIBITOR PROTEIN (P14.5) (UK114 ANTIGEN HOMOLOG)
  - >gi|1177435|emb|CAA64670| sapiens]
  - % Identity: 58.8
  - Alignment Length: 85
  - Location of Alignment in SEQ ID NO 460: from 8 to

45 92

## (B) Polypeptide Sequence

- 50
- Pat. Appln. SEQ ID NO 461
  - Ceres seq\_id 2048333
  - Location of start within SEQ ID NO 459: at 79 nt.

268

- (C) Nomination and Annotation of Domains within
- Predicted Polypeptide(s)
- Domain of unknown function
  - Location within SEQ ID NO 461: from 1 to 66 aa.

5

## (Dp) Related Amino Acid Sequences

- Alignment No. 307
- gi No. 5032215
- Description: ref|NP\_005827.1|pUK114| translational inhibitor protein p14.5
- >gi|1717975|sp|P52758|UK14\_HUMAN 14.5 KD TRANSLATIONAL INHIBITOR PROTEIN (P14.5) (UK114 ANTIGEN HOMOLOG)
- >gi|1177435|emb|CAA64670| sapiens]
- % Identity: 58.8
- Alignment Length: 85
- Location of Alignment in SEQ ID NO 461: from 1 to

66

## (B) Polypeptide Sequence

- 20
- Pat. Appln. SEQ ID NO 462
  - Ceres seq\_id 2048334
  - Location of start within SEQ ID NO 459: at 139 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- 25
- (Dp) Related Amino Acid Sequences
  - Alignment No. 308
  - gi No. 5032215
  - Description: ref|NP\_005827.1|pUK114| translational inhibitor protein p14.5
  - >gi|1717975|sp|P52758|UK14\_HUMAN 14.5 KD TRANSLATIONAL INHIBITOR PROTEIN (P14.5) (UK114 ANTIGEN HOMOLOG)
  - >gi|1177435|emb|CAA64670| sapiens]
  - % Identity: 58.8
  - Alignment Length: 85
  - Location of Alignment in SEQ ID NO 462: from 1 to

46

## Maximum Length Sequence:

Public Genomic DNA:

gi No: 4589412

## Predicted Exons:

INIT	37105 ...	37091	OCKHANG-CDS
INTR	36413 ...	36339	OCKHANG-CDS
INTR	36224 ...	36176	OCKHANG-CDS
INTR	35977 ...	35698	OCKHANG-CDS
INTR	35603 ...	35471	OCKHANG-CDS
TERM	35193 ...	35119	OCKHANG-CDS

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 463

- Ceres seq\_id 2048466 269

(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 464

- Ceres seq\_id 2048467  
- Location of start within SEQ ID NO 463: at 1 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences

- Alignment No. 309

- gi No. 3688432

- Description: (AJ011705) 40S ribosomal protein S10

[Lumbricus rubellus]  
- % Identity: 50.6

- Alignment Length: 164

- Location of Alignment in SEQ ID NO 464: from 30  
to 193

(3) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 465

- Ceres seq\_id 2048468

- Location of start within SEQ ID NO 463: at 4 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences

- Alignment No. 310

- gi No. 3688432

- Description: (AJ011705) 40S ribosomal protein S10

[Lumbricus rubellus]  
- % Identity: 50.6

- Alignment Length: 164

- Location of Alignment in SEQ ID NO 465: from 29  
to 192

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 466

- Ceres seq\_id 2048469

- Location of start within SEQ ID NO 463: at 88 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences

- Alignment No. 311

- gi No. 3688432

- Description: (AJ011705) 40S ribosomal protein S10  
[Lumbricus rubellus]  
- % Identity: 50.6

- Alignment Length: 164

270  
- Location of Alignment in SEQ ID NO 466: from 1 to  
164

Maximum Length Sequence:

Public Genomic DNA:

gi No: 4914400

Predicted Exons:

INTR 97410 ... 97833 OCKHMG-CDNA

10 SINGLE 97525 ... 97797 OCKHMG-CDS

gi No: 4938473

Predicted Exons:

SINGLE 9769 ... 10041 GENBANK

15 INTR 9654 ... 10078 OCKHMG-CDNA

INTR 10721 ... 10880 OCKHMG-CDNA

SINGLE 9769 ... 10041 OCKHMG-CDS

20 (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 467

- Ceres seq\_id 2050485

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 468

- Ceres seq\_id 2050486

- Location of start within SEQ ID NO 467: at 2 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences

- Alignment No. 312

- gi No. 2493089

- Description: ATP SYNTHASE F CHAIN, MITOCHONDRIAL  
[Lumbricus rubellus]  
- % Identity: 29.7

- Alignment Length: 77  
- Location of Alignment in SEQ ID NO 468: from 52  
to 125

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 469

- Ceres seq\_id 2050487

- Location of start within SEQ ID NO 467: at 116 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences

- Alignment No. 313

- gi No. 2493089

- Description: ATP SYNTHASE F CHAIN, MITOCHONDRIAL  
[Lumbricus rubellus]  
- % Identity: 29.7

- Alignment Length: 77

- 271  
- % Identity: 29.7  
- Alignment Length: 77  
- Location of Alignment in SEQ ID NO 469: from 14  
to 87
- 5  
Maximum Length Sequence:  
Public Genomic DNA:  
gi No: 4454022  
Predicted Exons:  
INIT 55210 ... 55024 GENBANK  
TERM 54477 ... 54209 GENBANK  
INTR 55270 ... 55024 OCKHAMG-CDNA  
INTR 54477 ... 54299 OCKHAMG-CDNA  
15  
INIT 55210 ... 55024 OCKHAMG-CDS  
TERM 54477 ... 54209 OCKHAMG-CDS  
20  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 470  
- Ceres seq\_id 2050708
- 25  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 471  
- Ceres seq\_id 2050709  
- Location of start within SEQ ID NO 470: at 61 nt.
- 30  
(C) Nomination and Annotation of Domains within  
(Dp) Related Amino Acid Sequences  
- Alignment No. 314  
- gi No. 2129641  
- Description: major latex protein type 1 -  
Arabidopsis thaliana >gi1107493|emb|CAA630261 (X91960) major  
latex protein type1 [Arabidopsis thaliana]  
- % Identity: 71.3  
- Alignment Length: 150  
- Location of Alignment in: SEQ ID NO 471: from 1 to  
40 150
- 45  
Maximum Length Sequence:  
Public Genomic DNA:  
gi No: 4914454  
Predicted Exons:  
INTR 42446 ... 41456 OCKHAMG-CDS  
INTR 40358 ... 40280 OCKHAMG-CDS  
INTR 39551 ... 39481 OCKHAMG-CDS  
INTR 39415 ... 39322 OCKHAMG-CDS  
INTR 39280 ... 39268 OCKHAMG-CDS  
INTR 38937 ... 38658 OCKHAMG-CDS  
50

- 272  
INTR 38575 ... 38455 OCKHAMG-CDS  
TERM 38300 ... 38220 OCKHAMG-CDS  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 472  
- Ceres seq\_id 2050901
- 5  
10  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 473  
- Ceres seq\_id 2050902  
- Location of start within SEQ ID NO 472: at 3 nt.
- 15  
(C) Nomination and Annotation of Domains within  
(Dp) Related Amino Acid Sequences  
- Alignment No. 315  
- gi No. 3688432  
- Description: (AJ011705) 40S ribosomal protein S10  
[Lumbricus rubellus]  
- % Identity: 53.1  
- Alignment Length: 147  
- Location of Alignment in SEQ ID NO 473: from 416  
to 560
- 20  
25  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 474  
- Ceres seq\_id 2050903  
- Location of start within SEQ ID NO 472: at 123 nt.
- 30  
(C) Nomination and Annotation of Domains within  
(Dp) Related Amino Acid Sequences  
- Alignment No. 316  
- gi No. 3688432  
- Description: (AJ011705) 40S ribosomal protein S10  
[Lumbricus rubellus]  
- % Identity: 53.1  
- Alignment Length: 147  
- Location of Alignment in SEQ ID NO 474: from 376  
to 520
- 35  
40  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 475  
- Ceres seq\_id 2050904  
- Location of start within SEQ ID NO 472: at 678 nt.  
- Location of Signal Peptide Cleavage Site within SEQ  
ID NO 475: at 39 aa.
- 45  
50  
(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

273

## (Dp) Related Amino Acid Sequences

- Alignment No. 317
- gi No. 3688432
- Description: (A01.705) 40S ribosomal protein S10
- % Identity: 53.1
- Alignment Length: 147
- Location of Alignment in SEQ ID NO 475: from 191 to 335

10

## Maximum Length Sequence:

## Public Genomic DNA:

- gi No: 4725940
- Predicted Exons:

SINGLE 96089 ... 95736 GENBANK

INTR 96156 ... 95622 OCKHAMG-CDNA

SINGLE 96083 ... 95736 OCKHAMG-CDS

- gi No: 5823567
- Predicted Exons:

SINGLE 96089 ... 95736 GENBANK

INTR 96156 ... 95622 OCKHAMG-CDNA

SINGLE 96083 ... 95736 OCKHAMG-CDS

25

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 476
- Ceres seq\_id 2051325

30

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 477
- Ceres seq\_id 2051326

35

- Location of start within SEQ ID NO 476: at 68 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 477: at 30 aa.

40

## (C) Nomenclature and Annotation of Domains within Predicted Polypeptide(s)

- Plant lipid transfer protein family
- Location within SEQ ID NO 477: from 34 to 115 aa.

45

## (Dp) Related Amino Acid Sequences

- Alignment No. 318
- gi No. 4725952
- Description: (A1049730) AIR1A-like protein
- % Identity: 100
- Alignment Length: 117

50

274

## - Location of Alignment in SEQ ID NO 477: from 1 to 117

117

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 478
- Ceres seq\_id 2051327
- Location of start within SEQ ID NO 476: at 74 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 478: at 28 aa.

10

## (C) Nomenclature and Annotation of Domains within Predicted Polypeptide(s)

- Plant lipid transfer protein family
- Location within SEQ ID NO 478: from 32 to 113 aa.

15

## (Dp) Related Amino Acid Sequences

- Alignment No. 319
- gi No. 4725952
- Description: (A1049730) AIR1A-like protein
- % Identity: 100
- Alignment Length: 117
- Location of Alignment in SEQ ID NO 478: from 1 to 115

20

## (Arabidopsis thaliana)

- % Identity: 100
- Alignment Length: 117
- Location of Alignment in SEQ ID NO 478: from 1 to 115

115

25

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 479
- Ceres seq\_id 2051328
- Location of start within SEQ ID NO 476: at 288 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 479: at 61 aa.

30

## (C) Nomenclature and Annotation of Domains within Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

35

- Maximum Length Sequence:
- Public Genomic DNA:
- gi No: 4490717

40

## Predicted Exons:

INIT	87492 ...	87378	GENBANK
INTR	87034 ...	86928	GENBANK
INTR	86683 ...	86641	GENBANK
TERM	86445 ...	86402	GENBANK

45

INTR	87555 ...	87378	OCKHAMG-CDNA
INTR	87034 ...	86928	OCKHAMG-CDNA
INTR	86683 ...	86641	OCKHAMG-CDNA
INTR	86445 ...	86257	OCKHAMG-CDNA
INIT	87492 ...	87378	OCKHAMG-CDS

50



275  
 INTR 87034 ... 86928 OCKHAMG-CDS  
 INTR 86683 ... 86641 OCKHAMG-CDS  
 TERM 86445 ... 86432 OCKHAMG-CDS

5 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 480  
 - Ceres seq\_id 2051633

10 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 481  
 - Ceres seq\_id 2051634  
 - Location of start within SEQ ID NO 480: at 1 nt.

15 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences

- Alignment No. 320  
 - gi No. 1083282  
 - Description: cytochrome-c oxidase (EC 1.9.3.1) -  
 mouse >gi1567766 (L06465) cytochrome c oxidase [Mus musculus]  
 >gi11094404|prf112106151A cytochrome c oxidase:SUBUNIT-VIA  
 [Mus musculus]  
 - % Identity: 41.4  
 - Alignment Length: 106  
 - Location of Alignment in SEQ ID NO 481: from 22  
 to 120

30 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 482  
 - Ceres seq\_id 2051635  
 - Location of start within SEQ ID NO 480: at 64 nt.

35 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences

- Alignment No. 321  
 - gi No. 1083282  
 - Description: cytochrome-c oxidase (EC 1.9.3.1) -  
 mouse >gi1567766 (L06465) cytochrome c oxidase [Mus musculus]  
 >gi11094404|prf112106151A cytochrome c oxidase:SUBUNIT-VIA  
 [Mus musculus]  
 - % Identity: 41.4  
 - Alignment Length: 106  
 - Location of Alignment in SEQ ID NO 482: from 1 to

45 99  
 Maximum Length Sequence:  
 Public Genomic DNA:  
 gi No: 4584531  
 Predicted Exons:

50

276  
 INTR 56873 ... 56641 OCKHAMG-CDNA  
 INTR 56627 ... 56451 OCKHAMG-CDNA  
 SINGLE 56812 ... 56645 OCKHAMG-CDS

5 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 483  
 - Ceres seq\_id 2051906

10 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 484  
 - Ceres seq\_id 2051907  
 - Location of start within SEQ ID NO 483: at 2 nt.

15 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences

- Alignment No. 322  
 - gi No. 4850407  
 - Description: (AC007357) EST gb|T21221 comes from  
 this gene. [Arabidopsis thaliana]  
 - % Identity: 57.1  
 - Alignment Length: 35  
 - Location of Alignment in SEQ ID NO 484: from 37  
 to 70

30 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 485  
 - Ceres seq\_id 2051908  
 - Location of start within SEQ ID NO 483: at 62 nt.

35 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences

- Alignment No. 323  
 - gi No. 4850407  
 - Description: (AC007357) EST gb|T21221 comes from  
 this gene. [Arabidopsis thaliana]  
 - % Identity: 57.1  
 - Alignment Length: 35  
 - Location of Alignment in SEQ ID NO 485: from 17  
 to 50

45 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 486  
 - Ceres seq\_id 2051909  
 - Location of start within SEQ ID NO 483: at 74 nt.

50 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)

277

## (Dp) Related Amino Acid Sequences

- Alignment No. 324
- gi No. 4850407
- Description: (AC007357) EST gb|T21221 comes from
- [Arabidopsis thaliana]
- % Identity: 57.1
- Alignment Length: 35
- Location of Alignment in SEQ ID NO 486: from 13 to 46

Maximum Length Sequence:  
Public Genomic DNA:

gi No: 4584841

## Predicted Exons:

- 15 INTR 75983 ... 75751 OCKHAMG-CDNA  
INTR 75412 ... 74936 OCKHAMG-CDNA
- 20 INTR 75992 ... 75751 OCKHAMG-CDNA  
INTR 75412 ... 75090 OCKHAMG-CDNA  
INIT 75937 ... 75751 OCKHAMG-CDS  
TERM 75412 ... 75132 OCKHAMG-CDS

## (Ac) cDNA Polynucleotide Sequence

- 25 - Pat. Appln. SEQ ID NO 487  
- Ceres seq\_id 2052403

## (B) Polypeptide Sequence

- 30 - Pat. Appln. SEQ ID NO 488  
- Ceres seq\_id 2052404  
- Location of start within SEQ ID NO 487: at 56 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- 35 - Pathogenesis-related protein Set V I family  
- Location within SEQ ID NO 488: from 5 to 155 aa.

## (Dp) Related Amino Acid Sequences

- 40 - Alignment No. 325
- gi No. 1321731
- Description: (Z72439) major allergen Cor a 1
- [Corylus avellana]
- % Identity: 36.3
- Alignment Length: 159
- Location of Alignment in SEQ ID NO 488: from 5 to 155

## Maximum Length Sequence:

Public Genomic DNA:

gi No: 4699904

278

## Predicted Exons:

- 5 INTR 82618 ... 82845 OCKHAMG-CDNA  
INTR 83226 ... 83301 OCKHAMG-CDNA  
INTR 83389 ... 83561 OCKHAMG-CDNA  
INTR 84124 ... 84385 OCKHAMG-CDNA
- 10 INIT 82730 ... 82845 OCKHAMG-CDS  
INTR 83226 ... 83301 OCKHAMG-CDS  
INTR 83389 ... 83561 OCKHAMG-CDS  
TERM 84124 ... 84199 OCKHAMG-CDS

## (Ac) cDNA Polynucleotide Sequence

- 15 - Pat. Appln. SEQ ID NO 489  
- Ceres seq\_id 2053545

## (B) Polypeptide Sequence

- 15 - Pat. Appln. SEQ ID NO 490  
- Ceres seq\_id 2053546
- 20 - Location of start within SEQ ID NO 489: at 113 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 490: at 36 aa.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- 25 - MAPEG family  
- Location within SEQ ID NO 490: from 59 to 136 aa.

## (Dp) Related Amino Acid Sequences

- 30 - gi No. 4758714
- Description: ref|NP\_004519.1|PMGST3| microsomal glutathione S-transferase 3 [gi|2583081 (AF026977) microsomal glutathione S-transferase 3 (Homo sapiens)]
- 35 - % Identity: 41.9
- Alignment Length: 131
- Location of Alignment in SEQ ID NO 490: from 8 to 136

## (B) Polypeptide Sequence

- 40 - Pat. Appln. SEQ ID NO 491
- Ceres seq\_id 2053547
- Location of start within SEQ ID NO 489: at 197 nt.
- (C) Nomination and Annotation of Domains within
- Predicted Polypeptide(s)
- MAPEG family
- Location within SEQ ID NO 491: from 31 to 108 aa.

## (Dp) Related Amino Acid Sequences

- 50 - Alignment No. 327

279

- gi No. 4758714  
- Description: ref|NP\_004519.1|pMGST3| microsomal  
glutathione S-transferase 3 >gi|2583081 (AF026977) microsomal  
glutathione S-transferase 3 (Homo sapiens)  
5 - % Identity: 41.9  
- Alignment Length: 131  
- Location of Alignment in SEQ ID NO 491: from 1 to 108

10 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 492  
- Ceres seq\_id 2053548  
- Location of start within SEQ ID NO 489: at 329 nt.  
- Location of Signal Peptide-Cleavage Site within SEQ  
15 ID NO 492: at 22 aa.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
- MAPEG family  
- Location within SEQ ID NO 492: from 1 to 64 aa.

20 (Dp) Related Amino Acid Sequences  
- Alignment No. 328  
- gi No. 4758714  
- Description: ref|NP\_004519.1|pMGST3| microsomal  
25 glutathione S-transferase 3 >gi|2583081 (AF026977) microsomal  
glutathione S-transferase 3 (Homo sapiens)  
- % Identity: 41.9  
- Alignment Length: 131  
30 - Location of Alignment in SEQ ID NO 492: from 1 to 64

Maximum Length Sequence:  
related to:

35 Clone IDs:  
2004  
31675  
266907  
Public Genomic DNA:  
40 gi No: 4887737  
Predicted Exons:  
INTR 7907 ... 8579 OCKHAMG-CDNA  
INTR 7913 ... 8587 OCKHAMG-CDNA  
45 INTR 7911 ... 8579 OCKHAMG-CDNA  
INTR 7909 ... 8579 OCKHAMG-CDNA  
50 INTR 7913 ... 8578 OCKHAMG-CDNA

INTR 7913 ... 8509 OCKHAMG-CDNA  
280  
SINGLE 7958 ... 8425 OCKHAMG-CDS  
gi No: 5019264  
5 Predicted Exons:  
SINGLE 4143 ... 4610 GENBANK  
INTR 4092 ... 4764 OCKHAMG-CDNA  
10 INTR 4098 ... 4772 OCKHAMG-CDNA  
INTR 4096 ... 4764 OCKHAMG-CDNA  
15 INTR 4094 ... 4764 OCKHAMG-CDNA  
INTR 4098 ... 4763 OCKHAMG-CDNA  
INTR 4098 ... 4694 OCKHAMG-CDNA  
20 SINGLE 4143 ... 4610 OCKHAMG-CDS  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 493  
- Ceres seq\_id 2053884  
- Alternative transcription start site(s) located in SEQ  
25 ID NO 493:  
-47,-3,2,3,4,5,6,7,8,10,11,12,13,14,15,16,20,21,28,29  
38,44,45,60,68  
- Clone 2004 starts at 1 and ends at 677 in cDNA.  
- Clone 31675 starts at 3 and ends at 607 in cDNA.  
30 - Clone 266907 starts at 4 and ends at 677 in cDNA.  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 494  
- Ceres seq\_id 2053885  
35 - Location of start within SEQ ID NO 493: at 1 nt.  
(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
40 - Alignment No. 329  
- gi No. 2498177  
- Description: BASP1 PROTEIN  
- % Identity: 28  
- Alignment Length: 164  
45 - Location of Alignment in SEQ ID NO 494: from 2 to 159  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 495  
50 - Ceres seq\_id 2053886  
- Location of start within SEQ ID NO 493: at 52 nt.

281

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 330
  - gi No. 2498177
  - Description: BASP1 PROTEIN
  - % Identity: 28
- Alignment Length: 164
- Location of Alignment: in SEQ ID NO 495: from 1 to 142

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 496
- Ceres seq\_id 2053887
- Location of start within SEQ ID NO 493: at 172 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 331
  - gi No. 2498177
  - Description: BASP1 PROTEIN
  - % Identity: 28
- Alignment Length: 164
- Location of Alignment: in SEQ ID NO 496: from 1 to 102

Maximum Length Sequence:

30 Public Genomic DNA:

gi No: 4887738

Predicted Exons:

- INTR 45004 ... 44725 OCKHAMG-CDNA
- INTR 44038 ... 43918 OCKHAMG-CDNA
- INTR 43839 ... 43444 OCKHAMG-CDNA

- INIT 44942 ... 44725 OCKHAMG-CDS
- INTR 44038 ... 43916 OCKHAMG-CDS
- TERM 43839 ... 43564 OCKHAMG-CDS

gi No: 5103850

Predicted Exons:

- INIT 44944 ... 44727 GENBANK
- INTR 44040 ... 43920 GENBANK
- TERM 43841 ... 43566 GENBANK

- INTR 45006 ... 44727 OCKHAMG-CDNA
- INTR 44040 ... 43920 OCKHAMG-CDNA
- INTR 43841 ... 43446 OCKHAMG-CDNA

- INIT 44944 ... 44727 OCKHAMG-CDS
- INTR 44040 ... 43920 OCKHAMG-CDS

282

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 497
- Ceres seq\_id 2053908

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 498
- Ceres seq\_id 2053909
- Location of start within SEQ ID NO 497: at 63 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 332
  - gi No. 2245000
  - Description: (297341) LEFT like protein
  - % Identity: 31.4
- Alignment Length: 156
- Location of Alignment: in SEQ ID NO 498: from 17 to 164

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 499
- Ceres seq\_id 2053910
- Location of start within SEQ ID NO 497: at 168 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 499: at 23 aa.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 333
  - gi No. 2245000
  - Description: (297341) LEFT like protein
  - % Identity: 31.4
- Alignment Length: 156
- Location of Alignment: in SEQ ID NO 499: from 1 to 129

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 500
- Ceres seq\_id 2053911
- Location of start within SEQ ID NO 497: at 216 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

283

- Alignment No. 334
- gi No. 2245000
- Description: (Z97341) LET1 like protein [Arabidopsis thaliana]
- % Identity: 31.4
- Alignment Length: 156
- Location of Alignment in SEQ ID NO 500: from 1 to

113

## 10 Maximum Length Sequence:

Public Genomic DNA:

gi No: 4335711

Predicted Exons:

INTR 46235 ... 46470 OCKHANG-CDNA  
INTR 46624 ... 46739 OCKHANG-CDNA  
INTR 47125 ... 47423 OCKHANG-CDNA  
  
INIT 46285 ... 46470 OCKHANG-CDS  
INTR 46624 ... 46739 OCKHANG-CDS  
TERM 47125 ... 47413 OCKHANG-CDS

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 501
- Ceres seq\_id 2056123

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 502
- Ceres seq\_id 2056124
- Location of start within SEQ ID NO 501: at 47 nt.

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Zinc finger, C3HC4 type (RING finger)
- Location within SEQ ID NO 502: from 150 to 190

aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 335
- gi No. 3790593
- Description: (AF079185) RING-H2 finger protein [Arabidopsis thaliana]
- % Identity: 55.1
- Alignment Length: 49
- Location of Alignment in SEQ ID NO 502: from 146

to 194

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 503
- Ceres seq\_id 2056125
- Location of start within SEQ ID NO 501: at 56 nt.

284

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Zinc finger, C3HC4 type (RING finger)
- Location within SEQ ID NO 503: from 147 to 197

aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 336
- gi No. 3790593
- Description: (AF079185) RING-H2 finger protein [Arabidopsis thaliana]
- % Identity: 55.1
- Alignment Length: 49
- Location of Alignment in SEQ ID NO 503: from 143

to 191

Maximum Length Sequence:

Public Genomic DNA:

gi No: 4432829

Predicted Exons:

INTR 51185 ... 51056 OCKHANG-CDNA  
INTR 50590 ... 50253 OCKHANG-CDNA  
  
INIT 56176 ... 55097 OCKHANG-CDS  
INTR 55014 ... 54161 OCKHANG-CDS  
INTR 54076 ... 54011 OCKHANG-CDS  
INTR 51530 ... 51382 OCKHANG-CDS  
INTR 51187 ... 51056 OCKHANG-CDS  
TERM 50590 ... 50418 OCKHANG-CDS

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 504
- Ceres seq\_id 2056245

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 505
- Ceres seq\_id 2056246
- Location of start within SEQ ID NO 504: at 1 nt.

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 337
- gi No. 4938484
- Description: (AL078464) transcription factor-like protein [Arabidopsis thaliana]
- % Identity: 56.2
- Alignment Length: 661

285

644 - Location of Alignment in SEQ ID NO 505: from 4 to

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 506
- Ceres seq\_id 2056247
- Location of start within SEQ ID NO 504: at 55 nt.

## (C) Nomination and Annotation of Domains within

## 10 Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 338
- gi No. 4938484
- Description: (AL078464) transcription factor-like protein [Arabidopsis thaliana]
- % Identity: 56.2
- Alignment Length: 661
- Location of Alignment in SEQ ID NO 506: from 1 to

626

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 507
- Ceres seq\_id 2056243
- Location of start within SEQ ID NO 504: at 751 nt.

## (C) Nomination and Annotation of Domains within

## 25 Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 339
- gi No. 4938484
- Description: (AL078464) transcription factor-like protein [Arabidopsis thaliana]
- % Identity: 56.2
- Alignment Length: 661
- Location of Alignment in SEQ ID NO 507: from 1 to

394

## Maximum Length Sequence:

## related to:

40 Clone IDs:

213322

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 508
- Ceres seq\_id 1941142

45

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 509
- Ceres seq\_id 1941143
- Location of start within SEQ ID NO 508: at 119 nt.

50

286

5 (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein L24e
- Location within SEQ ID NO 509: from 3 to 73 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 340
- gi No. 710521
- Description: 60S RIBOSOMAL PROTEIN L24
- >gl1154859|emb|CAA63960| (X94296) L24 ribosomal protein [Hordeum vulgare]
- % Identity: 94.4
- Alignment Length: 162
- Location of Alignment in SEQ ID NO 509: from 1 to

15 161

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 510
- Ceres seq\_id 1941144
- Location of start within SEQ ID NO 508: at 284 nt.

## (C) Nomination and Annotation of Domains within

## 25 Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 341
- gi No. 1710521
- Description: 60S RIBOSOMAL PROTEIN L24
- >gl1154859|emb|CAA63960| (X94296) L24 ribosomal protein [Hordeum vulgare]
- % Identity: 94.4
- Alignment Length: 162
- Location of Alignment in SEQ ID NO 510: from 1 to

30 106

## Maximum Length Sequence:

## related to:

35 Clone IDs:

241379

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 511
- Ceres seq\_id 1942975
- Alternative transcription start site(s) located in SEQ ID NO 511: 29, 30, 31, 32, 33, 36

45

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 512
- Ceres seq\_id 1942976
- Location of start within SEQ ID NO 511: at 1 nt.

50

287

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein S3, C-terminal domain.
- Location within SEQ ID NO 512: from 149 to 233

5 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 342
- gi No. 133940
- Description: 40S RIBOSOMAL PROTEIN S3A (S1A)
- >gi|70851|pir||R3XL3A ribosomal protein S3a - African clawed frog >gi|6509|emb|CAA40592| (X57322) ribosomal protein S1a [Xenopus laevis]
- % Identity: 82.5
- Alignment Length: 223
- Location of Alignment: in SEQ ID NO 512: from 46 to 268

15

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 513
- Ceres seq\_id 1942977
- Location of start within SEQ ID NO 511: at 136 nt.

20

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein S3, C-terminal domain.
- Location within SEQ ID NO 513: from 104 to 188

aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 343
- gi No. 133940
- Description: 40S RIBOSOMAL PROTEIN S3A (S1A)
- >gi|70851|pir||R3XL3A ribosomal protein S3a - African clawed frog >gi|6509|emb|CAA40592| (X57322) ribosomal protein S1a [Xenopus laevis]
- % Identity: 82.5
- Alignment Length: 223
- Location of Alignment in SEQ ID NO 513: from 1 to 223

40 223

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 514
- Ceres seq\_id 1942978
- Location of start within SEQ ID NO 511: at 205 nt.

45

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein S3, C-terminal domain.
- Location within SEQ ID NO 514: from 81 to 165 aa.

50

288

## (Dp) Related Amino Acid Sequences

- Alignment No. 344
- gi No. 133940
- Description: 40S RIBOSOMAL PROTEIN S3A (S1A)
- >gi|70851|pir||R3XL3A ribosomal protein S3a - African clawed frog >gi|6509|emb|CAA40592| (X57322) ribosomal protein S1a [Xenopus laevis]
- % Identity: 82.5
- Alignment Length: 223
- Location of Alignment in SEQ ID NO 514: from 1 to 200

10

## Maximum Length Sequence:

- related to:
- Clone IDs: 289536
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 515
- Ceres seq\_id 1944349

20

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 516
- Ceres seq\_id 1944350
- Location of start within SEQ ID NO 515: at 423 nt.

25

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein L24c
- Location within SEQ ID NO 516: from 3 to 73 aa.

30

## (Dp) Related Amino Acid Sequences

- Alignment No. 345
- gi No. 1710521
- Description: 60S RIBOSOMAL PROTEIN L24
- >gi|1154859|emb|CAA63960| (X54296) L24 ribosomal protein [Hordeum vulgare]
- % Identity: 94.4
- Alignment Length: 162
- Location of Alignment in SEQ ID NO 516: from 1 to 161

35

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 517
- Ceres seq\_id 1944351
- Location of start within SEQ ID NO 515: at 588 nt.

45

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- (Dp) Related Amino Acid Sequences
- Alignment No. 346

50

289

- gi No. 1710521
- Description: 60S RIBOSOMAL PROTEIN L24
- >gi1154859|emb|CAA639601 (X94296) L24 ribosomal protein [Hordium vulgare]
- % Identity: 94.4
- Alignment Length: 162
- Location of Alignment in SEQ ID NO 517: from 1 to 106

- 10 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 518
  - Ceres seq\_id 1944352
  - Location of start within SEQ ID NO 515: at 736 nt.

- 15 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 347
  - gi No. 1362587

- 20 precursor - longfin squid
- Description: spermatid-specific protein T2
  - % Identity: 55.8
  - Alignment Length: 52
  - Location of Alignment in SEQ ID NO 518: from 1 to 46

- 25 Maximum Length Sequence:  
related to:

- 30 Clone IDs:  
291258
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 519
  - Ceres seq\_id 1964011
  - Alternative transcription start site(s) located in SEQ ID NO 519: 4,34,37,38,39,40,41

- 35 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 520
  - Ceres seq\_id 1964012
  - Location of start within SEQ ID NO 519: at 3 nt.

- 40 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- Ribosomal protein S3, C-terminal domain.
  - Location within SEQ ID NO 520: from 150 to 234 aa.

- 50 (Dp) Related Amino Acid Sequences
- Alignment No. 348
  - gi No. 1173253

290

- Description: 40S RIBOSOMAL PROTEIN S3
- >gi1543317|pir||S41170 ribosomal protein S3 - mouse
- >gi157728|emb|CAA359161 (X51536) ribosomal protein S3 (AA 1-243) [Rattus rattus] >gi1439522|emb|CAA541671 (X76772) ribosomal protein S3 [Mus musculus]
- % Identity: 81.1
- Alignment Length: 227
- Location of Alignment in SEQ ID NO 520: from 47 to 273

- 10 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 521
  - Ceres seq\_id 1964013
  - Location of start within SEQ ID NO 519: at 141 nt.

- 15 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- Ribosomal protein S3, C-terminal domain.
  - Location within SEQ ID NO 521: from 104 to 163 aa.

- 20 (Dp) Related Amino Acid Sequences
- Alignment No. 349
  - gi No. 1173253

- 25 >gi1543317|pir||S41170 ribosomal protein S3 - mouse
- >gi157728|emb|CAA359161 (X51536) ribosomal protein S3 (AA 1-243) [Rattus rattus] >gi1439522|emb|CAA541671 (X76772) ribosomal protein S3 [Mus musculus]
- % Identity: 81.1
  - Alignment Length: 227
  - Location of Alignment in SEQ ID NO 521: from 1 to 227

- 30 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 522
  - Ceres seq\_id 1964014
  - Location of start within SEQ ID NO 519: at 210 nt.

- 35 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- Ribosomal protein S3, C-terminal domain.
  - Location within SEQ ID NO 522: from 81 to 165 aa.

- 40 (Dp) Related Amino Acid Sequences
- Alignment No. 350
  - gi No. 1173253

- 45 >gi1543317|pir||S41170 ribosomal protein S3 - mouse
- >gi157728|emb|CAA359161 (X51536) ribosomal protein S3 (AA 1-227)



291

243) [Rattus rattus] >gi1439522|emb|CAA541671 (X76772)  
ribosomal protein S3 [Mus musculus]

- % Identity: 81.1
- Alignment Length: 227
- Location of Alignment in SEQ ID NO 522: from 1 to

5 204

Maximum Length Sequence:  
related to:

1C Clone IDs:

207986  
224937  
227089  
229580  
241662  
275880  
278084

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 523
- Ceres seq\_id 1983854
- Alternative transcription start site(s) located in SEQ ID NO 523:

- 1,2,6,30,31,32,33,34,36
- Clone 207986 starts at 31 and ends at 810 in cDNA.
- Clone 224937 starts at 31 and ends at 810 in cDNA.
- Clone 227089 starts at 33 and ends at 804 in cDNA.
- Clone 229580 starts at 31 and ends at 821 in cDNA.
- Clone 241662 starts at 34 and ends at 815 in cDNA.
- Clone 275880 starts at 1 and ends at 804 in cDNA.
- Clone 278084 starts at 31 and ends at 815 in cDNA.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 524
- Ceres seq\_id 1983855
- Location of start within SEQ ID NO 523: at 315 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 351
- gi No. 1710521

>gi1154859|emb|CAA63960| (X94296) L24 ribosomal protein  
[Hordeum vulgare]

- % Identity: 94
- Alignment Length: 150
- Location of Alignment in SEQ ID NO 524: from 1 to

106

50 Maximum Length Sequence:  
related to:

292

Clone IDs:

221977  
226126  
293001

- 5 (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 525
- Ceres seq\_id 1990261
- Alternative transcription start site(s) located in SEQ ID NO 525:

- 10 16,30,32
- Clone 226126 starts at 30 and ends at 772 in cDNA.
- Clone 293001 starts at 32 and ends at 810 in cDNA.

(B) Polypeptide Sequence

- 15 - Pat. Appln. SEQ ID NO 526
- Ceres seq\_id 2061972
- Location of start within SEQ ID NO 525: at 144 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Ribosomal protein L24e
- Location within SEQ ID NO 526: from 3 to 73 aa.

(Dp) Related Amino Acid Sequences

- 25 - Alignment No. 352
- gi No. 1710521

>gi1154859|emb|CAA63960| (X94296) L24 ribosomal protein  
[Hordeum vulgare]

- 30 - % Identity: 94.4
- Alignment Length: 162
- Location of Alignment in SEQ ID NO 526: from 1 to

161

(B) Polypeptide Sequence

- 35 - Pat. Appln. SEQ ID NO 527
- Ceres seq\_id 2061973
- Location of start within SEQ ID NO 525: at 309 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 353
- gi No. 1710521

>gi1154859|emb|CAA63960| (X94296) L24 ribosomal protein  
[Hordeum vulgare]

- 45 - % Identity: 94.4
- Alignment Length: 162
- Location of Alignment in SEQ ID NO 527: from 1 to

106

293

**REF TABLE 2**

Maximum Length Sequence:

related to:

5 Clone IDs:

9458

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 1

- Ceres seq\_id 1007802

10 ID NO 1:  
- Alternative transcription start site(s) located in SEQ  
2,3,4

(B) Polypeptide Sequence

15 - Pat. Appln. SEQ ID NO 2

- Ceres seq\_id 1007803

- Location of start within SEQ ID NO 2: at 3 nt.

(B) Polypeptide Sequence

20 - Pat. Appln. SEQ ID NO 3

- Ceres seq\_id 1007804

- Location of start within SEQ ID NO 1: at 2 nt.

ID NO 3: at 29 aa.  
- Location of Signal Peptide Cleavage Site within SEQ

25 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 4

- Ceres seq\_id 1007805

- Location of start within SEQ ID NO 1: at 29 nt.

30 ID NO 4: at 20 aa.  
- Location of Signal Peptide Cleavage Site within SEQ

(Ba) Polypeptide Activities: Arabidopsis specific gene,

dicot

specific gene, plant

35 specific gene.

Maximum Length Sequence:

related to:

40 Clone IDs:

8458

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 5

- Ceres seq\_id 1008556

45 ID NO 5:  
- Alternative transcription start site(s) located in SEQ  
80,103,134,135,137,138,139,152,153,157,166,168

(B) Polypeptide Sequence

50 - Pat. Appln. SEQ ID NO 6

- Ceres seq\_id 1008557

294

- Location of start within SEQ ID NO 5: at 173 nt.

(Ba) Polypeptide Activities: Arabidopsis specific gene,

dicot

specific gene, plant

5 specific gene.

Maximum Length Sequence:

related to:

10 Clone IDs:

8384

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 7

- Ceres seq\_id 1008628

15 ID NO 7:  
- Alternative transcription start site(s) located in SEQ  
3,28

(B) Polypeptide Sequence

20 - Pat. Appln. SEQ ID NO 8

- Ceres seq\_id 1008629

- Location of start within SEQ ID NO 7: at 41 nt.

ID NO 7: at 41 nt.  
(Ba) Polypeptide Activities: Arabidopsis specific

25 gene, dicot

specific gene, plant

specific gene.

Maximum Length Sequence:

related to:

30 Clone IDs:

7260

Public Genomic DNA:

gi No: 3733889

35 Predicted Exons:

INIT 21286 ...

INTR 21606 ...

TERM 21951 ...

INIT 21286 ...

INTR 21606 ...

TERM 21951 ...

gi No: 6041810

Predicted Exons:

INIT 34915 ...

INTR 35235 ...

TERM 35580 ...

gi No: 6091711

Predicted Exons:

INIT 100255...

INTR 100575...

INIT 100478

INTR 100801

50 CCKHAMG-CDS  
CCKHAMG-CDS  
CCKHAMG-CDS  
CCKHAMG-CDS

- 295  
TERM 100920... 100927 OCKHAMG-CDS  
gi No: 6102641  
Predicted Exons:  
INIT 91941 ... 92164 OCKHAMG-CDS  
INTR 92261 ... 92487 OCKHAMG-CDS  
TERM 92636 ... 92613 OCKHAMG-CDS  
gi No: 6453849  
Predicted Exons:  
INIT 19658 ... 19435 OCKHAMG-CDS  
INTR 19338 ... 19112 OCKHAMG-CDS  
TERM 18993 ... 18986 OCKHAMG-CDS  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 9  
- Ceres seq\_id 1009376  
- Alternative transcription start site(s) located in SEQ ID NO 9: 4, 7, 22, 48, 56, 77
- 5  
20  
25  
30  
35  
40  
45  
50
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 10  
- Ceres seq\_id 1009377  
- Location of start within SEQ ID NO 9: at 238 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 10: at 76 aa.
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 11  
- Ceres seq\_id 1009378  
- Location of start within SEQ ID NO 9: at 334 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 11: at 44 aa.
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 12  
- Ceres seq\_id 1009379  
- Location of start within SEQ ID NO 9: at 245 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 12: at 47 aa.
- (Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot  
specific gene, plant
- Maximum Length Sequence:  
related to:  
Clone IDs:  
4489  
- Pat. Appln. SEQ ID NO 13
- (Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 13

- 296  
- Ceres seq\_id 1011128  
- Alternative transcription start site(s) located in SEQ ID NO 13: 6, 7
- 5  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 14  
- Ceres seq\_id 1011129  
- Location of start within SEQ ID NO 13: at 41 nt.
- 10  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 15  
- Ceres seq\_id 1011130  
- Location of start within SEQ ID NO 13: at 59 nt.
- 15  
(Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot  
specific gene, plant
- 20  
Maximum Length Sequence:  
related to:  
Clone IDs:  
42241
- 25  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 16  
- Ceres seq\_id 1011718
- 30  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 17  
- Ceres seq\_id 1011719  
- Location of start within SEQ ID NO 16: at 3 nt.
- 35  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 18  
- Ceres seq\_id 1011720  
- Location of start within SEQ ID NO 16: at 78 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 18: at 21 aa.
- 40  
(3) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 19  
- Ceres seq\_id 1011721  
- Location of start within SEQ ID NO 16: at 102 nt.
- 45  
(Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot  
specific gene, plant
- 50

297

Maximum Length Sequence:  
related to:

5 Clone IDs:  
42200

(Ac) cDNA Polynucleotide Sequence

10 - Pat. Appln. SEQ ID NO 20  
- Ceres seq\_id 1011735

(B) Polypeptide Sequence

15 - Pat. Appln. SEQ ID NO 21  
- Ceres seq\_id 1011736  
- Location of start within SEQ ID NO 20: at 75 nt.

(3) Polypeptide Sequence

20 - Pat. Appln. SEQ ID NO 22  
- Ceres seq\_id 1011737  
- Location of start within SEQ ID NO 20: at 1 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 22: at 37 aa.

(B) Polypeptide Sequence

25 - Pat. Appln. SEQ ID NO 23  
- Ceres seq\_id 1011738  
- Location of start within SEQ ID NO 20: at 2 nt.

(Ba) Polypeptide Activities: Arabidopsis specific gene,

30 dicot  
specific gene, plant

specific gene.

Maximum Length Sequence:

35 related to:  
Clone IDs:

42217

(Ac) cDNA Polynucleotide Sequence

40 - Pat. Appln. SEQ ID NO 24  
- Ceres seq\_id 1011755  
- Alternative transcription start site(s) located in SEQ ID NO 24:  
34

(B) Polypeptide Sequence

45 - Pat. Appln. SEQ ID NO 25  
- Ceres seq\_id 1011756  
- Location of start within SEQ ID NO 24: at 1 nt.

(B) Polypeptide Sequence

50 - Pat. Appln. SEQ ID NO 26

298

- Ceres seq\_id 1011757  
- Location of start within SEQ ID NO 24: at 19 nt.

(Ba) Polypeptide Activities: Arabidopsis specific

5 gene, dicot  
specific gene, plant

specific gene.

Maximum Length Sequence:

10 related to:  
Clone IDs:

42034

(Ac) cDNA Polynucleotide Sequence

15 - Pat. Appln. SEQ ID NO 27  
- Ceres seq\_id 1011032

(5) Polypeptide Sequence

20 - Pat. Appln. SEQ ID NO 28  
- Ceres seq\_id 1011033  
- Location of start within SEQ ID NO 27: at 97 nt.

(B) Polypeptide Sequence

25 - Pat. Appln. SEQ ID NO 29  
- Ceres seq\_id 1011034  
- Location of start within SEQ ID NO 27: at 100 nt.

(B) Polypeptide Sequence

30 - Pat. Appln. SEQ ID NO 30  
- Ceres seq\_id 1011035  
- Location of start within SEQ ID NO 27: at 106 nt.

(Ba) Polypeptide Activities: Arabidopsis specific gene,

35 dicot  
specific gene, plant

specific gene.

Maximum Length Sequence:

40 related to:  
Clone IDs:

41812

Public Genomic DNA:

gi No: 6434227

2-predicted Exons:

45 INIT 65854 ... 65974 OCKHAMG-CDS  
TERM 66067 ... 66206 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 31  
- Ceres seq\_id 1011907

(B) Polypeptide Sequence

50

- 299
- 5 ID NO 32: at 30 aa.
- Pat. Appln. SEQ ID NO 32
  - Ceres seq\_id 1011908
  - Location of start within SEQ ID NO 31: at 218 nt.
  - Location of Signal Peptide Cleavage Site within SEQ

## (B) Polypeptide Sequence

- 10
- Pat. Appln. SEQ ID NO 33
  - Ceres seq\_id 1011909
  - Location of start within SEQ ID NO 31: at 239 nt.
  - Location of Signal Peptide Cleavage Site within SEQ
- ID NO 33: at 23 aa.

## (B) Polypeptide Sequence

- 15
- Pat. Appln. SEQ ID NO 34
  - Ceres seq\_id 1011910
  - Location of start within SEQ ID NO 31: at 3 nt.
  - Location of Signal Peptide Cleavage Site within SEQ
- ID NO 34: at 17 aa.

20 (Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot

specific gene, plant

## Maximum Length Sequence:

related to:  
Clone IDs:

41828

30 Public Genomic DNA:

gi No: 4159707

Predicted Exons:

- 35
- INTR 15084 ... 15201 OCKHAMG-CDNA
  - INTR 15295 ... 15324 OCKHAMG-CDNA
  - INTR 15414 ... 15533 OCKHAMG-CDNA
  - INTR 15648 ... 15956 OCKHAMG-CDNA

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 35
- Ceres seq\_id 1011911

## (B) Polypeptide Sequence

- 40
- Pat. Appln. SEQ ID NO 36
  - Ceres seq\_id 1011912
  - Location of start within SEQ ID NO 35: at 194 nt.

## (B) Polypeptide Sequence

- 45
- Pat. Appln. SEQ ID NO 37
  - Ceres seq\_id 1011913
  - Location of start within SEQ ID NO 35: at 230 nt.
  - Location of Signal Peptide Cleavage Site within SEQ
- ID NO 37: at 22 aa.

300

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 38
  - Ceres seq\_id 1011914
  - Location of start within SEQ ID NO 35: at 2 nt.
  - Location of Signal Peptide Cleavage Site within SEQ
- ID NO 38: at 28 aa.

10 (Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot

specific gene, plant

## Maximum Length Sequence:

related to:

Clone IDs:

41723

## (Ac) cDNA Polynucleotide Sequence

- 20
- Pat. Appln. SEQ ID NO 39
  - Ceres seq\_id 1011954

## (B) Polypeptide Sequence

- 25
- Pat. Appln. SEQ ID NO 40
  - Ceres seq\_id 1011955
  - Location of start within SEQ ID NO 39: at 1 nt.
  - Location of Signal Peptide Cleavage Site within SEQ
- ID NO 40: at 40 aa.

## (B) Polypeptide Sequence

- 30
- Pat. Appln. SEQ ID NO 41
  - Ceres seq\_id 1011956
  - Location of start within SEQ ID NO 39: at 107 nt.

35 (Ba) Polypeptide Activities: Similar to DNAB Protein  
Activities

## Maximum Length Sequence:

related to:

Clone IDs:

41730

## Public Genomic DNA:

gi No: 4678371

Predicted Exons:

SINGLE 48996 ... 49184

GENBANK

SINGLE 48996 ... 49184

OCKHAMG-CDS

## (Ac) cDNA Polynucleotide Sequence

- 50
- Pat. Appln. SEQ ID NO 42
  - Ceres seq\_id 1011960

301  
- Alternative transcription start site(s) located in SEQ ID NO 42:  
20, 22, 25, 43

5 (B) Polypeptide Sequence:

- Pat. Appln. SEQ ID NO 43
- Ceres seq\_id 1011961
- Location of start within SEQ ID NO 42: at 69 nt.

10 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 44
- Ceres seq\_id 1011962
- Location of start within SEQ ID NO 42: at 78 nt.

15 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 45
- Ceres seq\_id 1011963
- Location of start within SEQ ID NO 42: at 90 nt.

20 (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot

specific gene, plant specific gene, similar to serine rich protein activities.

25 Maximum Length Sequence:

related to:  
Clone IDs:

38977

30 (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 46
- Ceres seq\_id 1014075
- Alternative transcription start site(s) located in SEQ ID NO 46:  
-3, -2, -1, 2, 3, 4, 5, 6, 7, 8, 10, 15, 103

35 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 47
- Ceres seq\_id 1014076
- Location of start within SEQ ID NO 46: at 1 nt.

40 (D) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 48
- Ceres seq\_id 1014077
- Location of start within SEQ ID NO 46: at 84 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 48: at 22 aa.

50 (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot

302  
specific gene.  
specific gene, plant

5 Maximum Length Sequence:

related to:

Clone IDs:

37081

Public Genomic DNA:

gi No: 4159707

Predicted Exons:

INIT 15458 ... 15533 OCKHANG-CDS

TERM 15648 ... 15751 OCKHANG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 49
- Ceres seq\_id 1015865
- Alternative transcription start site(s) located in SEQ ID NO 49:  
-29

20 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 50
- Ceres seq\_id 1015866
- Location of start within SEQ ID NO 49: at 281 nt.

25 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 51
- Ceres seq\_id 1015867
- Location of start within SEQ ID NO 49: at 317 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 51: at 22 aa.

30 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 52
- Ceres seq\_id 1015868
- Location of start within SEQ ID NO 49: at 3 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 52: at 15 aa.

40 (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot

specific gene.  
specific gene, plant

45 Maximum Length Sequence:

related to:

Clone IDs:

29375

Public Genomic DNA:

gi No: 4159707

Predicted Exons:

INTR 15158 ... 15201 OCKHANG-CDNA

303  
INTR 15449 ... 15533 OCKHAMG-CDNA  
INTR 15648 ... 15947 OCKHAMG-CDNA

(Ac) cDNA Polynucleotide Sequence

- 5 - Pat. Appln. SEQ ID NO 53  
- Ceres seq\_id 1021371  
- Alternative transcription start site(s) located in SEQ ID NO 53:  
-74

10 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 54  
- Ceres seq\_id 1021372  
- Location of start within SEQ ID NO 53: at 3 nt.

15 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 55  
- Ceres seq\_id 102-373  
- Location of start within SEQ ID NO 53: at 54 nt.

20 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 56  
- Ceres seq\_id 1021374  
- Location of start within SEQ ID NO 53: at 90 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 56: at 22 aa.

(Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot

30 specific gene, plant

Maximum Length Sequence:

related to:

Clone IDs:

- 35 27064  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 57  
- Ceres seq\_id 1022578

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 58  
- Ceres seq\_id 1022579  
- Location of start within SEQ ID NO 57: at 1 nt.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 59  
- Ceres seq\_id 1022580  
- Location of start within SEQ ID NO 57: at 157 nt.

(B) Polypeptide Sequence

304

- Pat. Appln. SEQ ID NO 60  
- Ceres seq\_id 1022581  
- Location of start within SEQ ID NO 57: at 10 nt.

5 (Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot

specific gene, plant

10 Maximum Length Sequence:

related to:

Clone IDs:

23773

Public Genomic DNA:

gi No: 6136357

Predicted Exons:

- INTR 51972 ... 51830 OCKHAMG-CDNA  
INTR 51129 ... 51055 OCKHAMG-CDNA  
INTR 50948 ... 50806 OCKHAMG-CDNA  
INTR 50706 ... 50608 OCKHAMG-CDNA  
INTR 50326 ... 50147 OCKHAMG-CDNA

gi No: 6143855

Predicted Exons:

- INTR 51972 ... 51830 OCKHAMG-CDNA  
INTR 51129 ... 51055 OCKHAMG-CDNA  
INTR 50948 ... 50806 OCKHAMG-CDNA  
INTR 50706 ... 50608 OCKHAMG-CDNA  
INTR 50326 ... 50147 OCKHAMG-CDNA

(Ac) cDNA Polynucleotide Sequence

- 30 - Pat. Appln. SEQ ID NO 61  
- Ceres seq\_id 1024240

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 62  
- Ceres seq\_id 1024241  
- Location of start within SEQ ID NO 61: at 54 nt.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 63  
- Ceres seq\_id 1024242  
- Location of start within SEQ ID NO 61: at 66 nt.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 64  
- Ceres seq\_id 1024243  
- Location of start within SEQ ID NO 61: at 404 nt.

(Ba) Polypeptide Activities: Arabidopsis specific  
gene, dicot

specific gene, plant

305

Maximum Length Sequence:  
related to:

Clone IDs:  
2030

Public Genomic DNA:

gi No: 5441914

Predicted Exons:

SINGLE 4802 ... 4449 OCKHAMG-CDS

gi No: 5881519

Predicted Exons:

SINGLE 4802 ... 4449 GENBANK

SINGLE 4802 ... 4449 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 65

- Ceres seq\_id 1026562

- Alternative transcription start site(s) located in SEQ

ID NO 65:

36,41,43,44,88,90,94,102

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 66

- Ceres seq\_id 1026563

- Location of start within SEQ ID NO 65: at 1 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 66: at 51 aa.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 67

- Ceres seq\_id 1026564

- Location of start within SEQ ID NO 65: at 109 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 67: at 15 aa.

(Ba) Polypeptide Activities: Similar to cytochrome c  
oxidase II activities.

Maximum Length Sequence:  
related to:

Clone IDs:  
20097

Public Genomic DNA:

gi No: 5733893

Predicted Exons:

INTR 66948 ... 67429 OCKHAMG-CDNA

SINGLE 66986 ... 67324 OCKHAMG-CDS

gi No: 5748882

Predicted Exons:

306

INTR 41416 ... 40935 OCKHAMG-CDNA

SINGLE 41380 ... 41042 OCKHAMG-CDS

gi No: 5801671

Predicted Exons:

INTR 51184 ... 50703 OCKHAMG-CDNA

SINGLE 51148 ... 50810 OCKHAMG-CDS

gi No: 5809708

Predicted Exons:

INTR 41416 ... 40935 OCKHAMG-CDNA

SINGLE 41380 ... 41042 OCKHAMG-CDS

gi No: 5836114

Predicted Exons:

INTR 42487 ... 42006 OCKHAMG-CDNA

SINGLE 42451 ... 42113 OCKHAMG-CDS

gi No: 5923662

Predicted Exons:

INTR 43106 ... 43587 OCKHAMG-CDNA

SINGLE 43144 ... 43482 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 68

- Ceres seq\_id 1026648

- Alternative transcription start site(s) located in SEQ

ID NO 68:

-74, -72, -48, -42, 76

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 69

- Ceres seq\_id 1026649

- Location of start within SEQ ID NO 68: at 39 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 69: at 59 aa.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 70

- Ceres seq\_id 1026650

- Location of start within SEQ ID NO 68: at 78 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 70: at 46 aa.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 71

- Ceres seq\_id 1026651

- Location of start within SEQ ID NO 68: at 81 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 71: at 45 aa.



307

(Ba) Polypeptide Activities: Arabidopsis specific

gene, dicot

specific gene, plant

specific gene.

5

Maximum Length Sequence:

related to:

Clone IDs:

18274

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 72

- Ceres seq\_id 1027881

- Alternative transcription start site(s) located in SEQ

ID NO 72:

-8,-4,2

15

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 73

- Ceres seq\_id 1027882

- Location of start within SEQ ID NO 72: at 3 nt.

20

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 74

- Ceres seq\_id 1027883

- Location of start within SEQ ID NO 72: at 36 nt.

25

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 75

- Ceres seq\_id 1027884

- Location of start within SEQ ID NO 72: at 105 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 75: at 18 aa.

30

(Ba) Polypeptide Activities: Arabidopsis specific

gene, dicot

specific gene, plant

specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

121894

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 76

- Ceres seq\_id 1381797

- Alternative transcription start site(s) located in SEQ

ID NO 76:

-4,-18,19,22,26

40

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 77

50

308

- Ceres seq\_id 1381798

- Location of start within SEQ ID NO 76: at 70 nt.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 78

- Ceres seq\_id 1381799

- Location of start within SEQ ID NO 76: at 82 nt.

5

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 79

- Ceres seq\_id 1381800

- Location of start within SEQ ID NO 76: at 109 nt.

10

(Ba) Polypeptide Activities: Arabidopsis specific

gene, dicot

specific gene, plant

specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

31667

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 80

- Ceres seq\_id 1442747

- Alternative transcription start site(s) located in SEQ

ID NO 80:

3,5,6,11,62,64,65,66,69,70,71,72,74

25

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 81

- Ceres seq\_id 1442748

- Location of start within SEQ ID NO 80: at 2 nt.

30

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 82

- Ceres seq\_id 1442749

- Location of start within SEQ ID NO 80: at 140 nt.

35

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 83

- Ceres seq\_id 1442750

- Location of start within SEQ ID NO 80: at 224 nt.

40

(Ba) Polypeptide Activities: Arabidopsis specific

gene, dicot

specific gene, plant

specific gene.

Maximum Length Sequence:

related to:

50

309

Clone IDs:  
270354

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 84
- Ceres seq\_id 1459199

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 85
- Ceres seq\_id 1459200
- Location of start within SEQ ID NO 84: at 2 nt.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 86
- Ceres seq\_id 1459201
- Location of start within SEQ ID NO 84: at 56 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 86: at 20 aa.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 87
- Ceres seq\_id 1459202
- Location of start within SEQ ID NO 84: at 80 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 87: at 17 aa.

(Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot

specific gene, plant

specific gene.

35 Maximum Length Sequence:  
related to:

Clone IDs:  
12250

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 88
- Ceres seq\_id 1565605
- Alternative transcription start site(s) located in SEQ ID NO 88:  
-17

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 89
- Ceres seq\_id 1565606
- Location of start within SEQ ID NO 88: at 133 nt.

(3) Polypeptide Sequence

310

- Pat. Appln. SEQ ID NO 90
- Ceres seq\_id 1565607
- Location of start within SEQ ID NO 88: at 181 nt.

(Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot

specific gene, plant

specific gene.

10 Maximum Length Sequence:  
related to:

Clone IDs:  
97883

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 91
- Ceres seq\_id 1566686
- Alternative transcription start site(s) located in SEQ ID NO 91:  
58, 61

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 92
- Ceres seq\_id 1566687
- Location of start within SEQ ID NO 91: at 137 nt.

(Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot

specific gene, plant

specific gene.

30 Maximum Length Sequence:  
related to:

Clone IDs:  
12408

35 (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 93
- Ceres seq\_id 1567367
- Alternative transcription start site(s) located in SEQ ID NO 93:  
-39, 6, 11, 14, -7, 25, 32, 33, 34, 35, 60

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 94
- Ceres seq\_id 1567368
- Location of start within SEQ ID NO 93: at 1 nt.

(3) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 95
- Ceres seq\_id 1567369
- Location of start within SEQ ID NO 93: at 2 nt.

311

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 96
- Ceres seq\_id 1567370
- Location of start within SEQ ID NO 93: at 65 nt.

5

(Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot

specific gene.

10

Maximum Length Sequence:

related to:

Clone IDs:

19528

Public Genomic DNA:

gi No: 6091770

Predicted Exons:

INIT 21323 ... 21390 OCKHAMG-CDS  
INTR 21467 ... 21564 OCKHAMG-CDS  
INTR 21995 ... 22027 OCKHAMG-CDS  
TERM 22386 ... 22468 OCKHAMG-CDS

20

gi No: 6102636

Predicted Exons:

INIT 21323 ... 21390 OCKHAMG-CDS  
INTR 21467 ... 21564 OCKHAMG-CDS  
INTR 21995 ... 22027 OCKHAMG-CDS  
TERM 22386 ... 22468 OCKHAMG-CDS

25

gi No: 6137907

Predicted Exons:

INIT 21252 ... 21319 OCKHAMG-CDS  
INTR 21396 ... 21493 OCKHAMG-CDS  
INTR 21924 ... 21956 OCKHAMG-CDS  
TERM 22315 ... 22397 OCKHAMG-CDS

30

gi No: 6437519

Predicted Exons:

INIT 14107 ... 14174 OCKHAMG-CDS  
INTR 14251 ... 14348 OCKHAMG-CDS  
INTR 14779 ... 14811 OCKHAMG-CDS  
TERM 15170 ... 15252 OCKHAMG-CDS

35

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 97
- Ceres seq\_id 1570101
- Alternative transcription start site(s) located in SEQ

ID NO 97:

9,29,30,31,32,39,40,43,44,61,62,64,65,198

45

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 98
- Ceres seq\_id 1570102
- Location of start within SEQ ID NO 97: at 116 nt.

50

312

(Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot

specific gene.

5

Maximum Length Sequence:

related to:

Clone IDs:

42395

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 99
- Ceres seq\_id 1571051

10

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 100
- Ceres seq\_id 1571052
- Location of start within SEQ ID NO 99: at 1 nt.

15

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 101
- Ceres seq\_id 1571053
- Location of start within SEQ ID NO 99: at 16 nt.
- Location of Signal Peptide Cleavage Site within SEQ

20

ID NO 101: at 30 aa.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 102
- Ceres seq\_id 1571054
- Location of start within SEQ ID NO 99: at 139 nt.

30

(Ba) Polypeptide Activities: Arabidopsis specific  
gene, dicot

specific gene.

35

Maximum Length Sequence:

related to:

Clone IDs:

6487

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 103
- Ceres seq\_id 1571100
- Alternative transcription start site(s) located in SEQ

ID NO 103:

5,6

45

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 104
- Ceres seq\_id 1571101
- Location of start within SEQ ID NO 103: at 1 nt.

50

- 313
- Location of Signal Peptide Cleavage Site within SEQ ID NO 104: at 19 aa.
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 105
  - Ceres seq\_id 1571102
  - Location of start within SEQ ID NO 103: at 82 nt.
- (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot
- specific gene, plant
- 10
- Maximum Length Sequence:  
related to:
- 15
- Clone IDs:
- 39977
  - 14890
  - 34623
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 106
  - Ceres seq\_id 1665272
  - Alternative transcription start site(s) located in SEQ ID NO 106:
- 25
- 3,15,32,33,34,35,37,38,39,40,41,42,43,44,45,46,47,48,49,50
- 51,60,61,63,66,69,70,71,72,74,75,76,78,79,84,85
- Clone 39977 starts at 35 and ends at 636 in cDNA.
  - Clone 14890 starts at 43 and ends at - in cDNA.
  - Clone 34623 starts at 1 and ends at 598 in cDNA.
- 30
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 107
  - Ceres seq\_id 1665273
  - Location of start within SEQ ID NO 106: at 90 nt.
- 35
- (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot
- specific gene, plant
- 40
- Maximum Length Sequence:  
related to:
- Clone IDs:
- 24562
- 45
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 108
  - Ceres seq\_id 1713895
  - Alternative transcription start site(s) located in SEQ ID NO 108:  
-7,2,3,5,7,8,13,17,29,30,32,34
- 50

- 314
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 109
  - Ceres seq\_id 1713896
  - Location of start within SEQ ID NO 108: at 3 nt.
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 110
  - Ceres seq\_id 1713397
  - Location of start within SEQ ID NO 108: at 57 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 110: at 20 aa.
- 10
- (3) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 111
  - Ceres seq\_id 1713898
  - Location of start within SEQ ID NO 108: at 81 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 111: at 17 aa.
- 20
- (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot
- specific gene, plant
- 15
- Maximum Length Sequence:  
related to:
- Clone IDs:
- 41879
- 30
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 112
  - Ceres seq\_id 1923752
- 35
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 113
  - Ceres seq\_id 1923753
  - Location of start within SEQ ID NO 112: at 2 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 113: at 19 aa.
- 40
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 114
  - Ceres seq\_id 1923754
  - Location of start within SEQ ID NO 112: at 292 nt.
- 45
- (3) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 115
  - Ceres seq\_id 1923755
  - Location of start within SEQ ID NO 112: at 384 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 115: at 27 aa.
- 50

315

(Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot  
specific gene, plant

5

Maximum Length Sequence:  
related to:  
Clone IDs:

i50069

- (Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 116  
- Ceres seq\_id 1976816  
- Alternative transcription start site(s) located in SEQ

ID NO 116:

20,30,32

15

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 117  
- Ceres seq\_id 1976817  
- Location of start within SEQ ID NO 116: at 2 nt.  
- Location of Signal Peptide Cleavage Site within SEQ

ID NO 117: at 61 aa.

20

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 118  
- Ceres seq\_id 1976818  
- Location of start within SEQ ID NO 116: at 3 nt.

25

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 119  
- Ceres seq\_id 1976819  
- Location of start within SEQ ID NO 116: at 52 nt.

30

(Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot  
specific gene, plant

35

Maximum Length Sequence:  
related to:  
Clone IDs:

109997

Public Genomic DNA:

gi No: 4263813

Predicted Exons:

INIT 82737 ... 82740 OCKHAMG-CDS  
TERM 83475 ... 83623 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 120  
- Ceres seq\_id 2025128

50

316

- Alternative transcription start site(s) located in SEQ  
ID NO 120:  
- 5,5,61,88  
- Clone 109997 starts at 5 and ends at 566 in cDNA.

5

(3) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 121  
- Ceres seq\_id 2025129  
- Location of start within SEQ ID NO 120: at 1 nt.

10

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 122  
- Ceres seq\_id 2025130  
- Location of start within SEQ ID NO 120: at 3 nt.  
- Location of Signal Peptide Cleavage Site within SEQ  
ID NO 122: at 21 aa.

15

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 123  
- Ceres seq\_id 2025131  
- Location of start within SEQ ID NO 120: at 316 nt.

20

(Ba) Polypeptide Activities: Arabidopsis specific  
gene, dicot  
specific gene, plant

25

Maximum Length Sequence:  
related to:

Clone IDs:

41792

Public Genomic DNA:

gi No: 5762549

Predicted Exons:

INTR 41272 ... 41749 OCKHAMG-CDNA

gi No: 5776585

Predicted Exons:

INTR 55175 ... 55652 OCKHAMG-CDNA

gi No: 5801669

Predicted Exons:

INTR 97802 ... 98279 OCKHAMG-CDNA

gi No: 5809709

Predicted Exons:

INTR 97803 ... 98280 OCKHAMG-CDNA

gi No: 5932531

Predicted Exons:

INTR 90652 ... 91129 OCKHAMG-CDNA

gi No: 6013612

Predicted Exons:

INTR 9143 ... 9620 OCKHAMG-CDNA

gi No: 6016671

50

317

Predicted Exons:  
INTR 9256 ... 9743 OCKHAMG-CDNA  
gi No: 6041828  
Predicted Exons:  
INTR 9266 ... 9743 OCKHAMG-CDNA  
gi No: 6175159  
Predicted Exons:  
INTR 92644 ... 92167 OCKHAMG-CDNA  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 124  
- Ceres seq\_id 2025402

10

(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 125  
- Ceres seq\_id 2025403  
- Location of start within SEQ ID NO 124: at 186 nt.

15

(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 126  
- Ceres seq\_id 2025404  
- Location of start within SEQ ID NO 124: at 2 nt.

20

(Ba) Polypeptide Activities: Similar to ribose 5-phosphate isomerase

25

Maximum Length Sequence:  
related to:  
Clone IDs:  
28572  
Public Genomic DNA:  
gi No: 4725940  
Predicted Exons:  
INTR 99197 ... 99167 OCKHAMG-CDS  
INTR 99060 ... 98960 OCKHAMG-CDS  
INTR 98863 ... 98753 OCKHAMG-CDS  
INTR 98654 ... 97823 OCKHAMG-CDS  
INTR 97730 ... 97477 OCKHAMG-CDS  
INTR 96772 ... 96683 OCKHAMG-CDS  
TERM 96157 ... 96086 OCKHAMG-CDS  
gi No: 5823567  
Predicted Exons:  
INTR 99197 ... 99167 OCKHAMG-CDS  
INTR 99060 ... 98960 OCKHAMG-CDS  
INTR 98863 ... 98753 OCKHAMG-CDS  
INTR 98654 ... 97823 OCKHAMG-CDS  
INTR 97730 ... 97477 OCKHAMG-CDS  
INTR 96772 ... 96683 OCKHAMG-CDS  
TERM 96157 ... 96086 OCKHAMG-CDS

30

(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 127  
- Ceres seq\_id 2025479

35

40

45

50

318

- Alternative transcription start site(s) located in SEQ ID NO 127:  
1419  
- Clone 28572 starts at 1419 and ends at 1955 in cDNA.

5

(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 128  
- Ceres seq\_id 2025480  
- Location of start within SEQ ID NO 127: at 1 nt.

10

(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 129  
- Ceres seq\_id 2025481  
- Location of start within SEQ ID NO 127: at 226 nt.

15

(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 130  
- Ceres seq\_id 2025482  
- Location of start within SEQ ID NO 127: at 271 nt.

20

(Ba) Polypeptide Activities: Similar Cell Wall Plasma Membrane disconnecting CLCF Protein activities, and proline rich protein activities.

25

Maximum Length Sequence:  
Public Genomic DNA:  
gi No: 6134247  
Predicted Exons:  
INTR 46663 ... 46055 OCKHAMG-CDNA  
SINGLE 46579 ... 46103 OCKHAMG-CDS

30

(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 131  
- Ceres seq\_id 2032963

35

(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 132  
- Ceres seq\_id 2032964  
- Location of start within SEQ ID NO 131: at 85 nt.

40

(Ba) Polypeptide Activities: Arabidopsis specific gene,  
elcot  
specific gene, plant

45

Maximum Length Sequence:  
Public Genomic DNA:  
gi No: 6449509  
Predicted Exons:

50

319  
 INIT 32110 ... 32118 OCKHAMG-CDS  
 INTR 32363 ... 32538 OCKHAMG-CDS  
 INTR 33006 ... 33084 OCKHAMG-CDS  
 TERM 33182 ... 33247 OCKHAMG-CDS

gi No: 6453848  
 Predicted Exons:

INIT 34795 ... 34803 OCKHAMG-CDS  
 INTR 35048 ... 35223 OCKHAMG-CDS  
 INTR 35691 ... 35769 OCKHAMG-CDS  
 TERM 35867 ... 35932 OCKHAMG-CDS

gi No: 6456153  
 Predicted Exons:

INIT 32305 ... 32313 OCKHAMG-CDS  
 INTR 32558 ... 32733 OCKHAMG-CDS  
 INTR 33201 ... 33279 OCKHAMG-CDS  
 TERM 33377 ... 33442 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 133  
 - Ceres seq\_id 2033706

(B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 134  
 - Ceres seq\_id 2033707  
 - Location of start within SEQ ID NO 133: at 1 nt.

(B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 135  
 - Ceres seq\_id 2033708  
 - Location of start within SEQ ID NO 133: at 34 nt.

(B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 136  
 - Ceres seq\_id 2033709  
 - Location of start within SEQ ID NO 133: at 160 nt.

(Ba) Polypeptide Activities: Arabidopsis specific gene,  
 dicot  
 specific gene, plant

Maximum Length Sequence:  
 Public Genomic DNA:

gi No: 4512646  
 Predicted Exons:  
 INTR 43130 ... 43387 OCKHAMG-CDNA  
 INTR 43479 ... 43638 OCKHAMG-CDNA

INIT 43224 ... 43387 OCKHAMG-CDS  
 TERM 43479 ... 43494 OCKHAMG-CDS

320  
 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 137  
 - Ceres seq\_id 2043118

(B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 138  
 - Ceres seq\_id 2043119  
 - Location of start within SEQ ID NO 137: at 65 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 138: at 22 aa.

(B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 139  
 - Ceres seq\_id 2043120  
 - Location of start within SEQ ID NO 137: at 1 nt.

(B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 140  
 - Ceres seq\_id 2043121  
 - Location of start within SEQ ID NO 137: at 95 nt.

(Ba) Polypeptide Activities: Arabidopsis specific gene,  
 dicot  
 specific gene, plant

Maximum Length Sequence:  
 Public Genomic DNA:

gi No: 3985934  
 Predicted Exons:  
 INTR 40094 ... 40516 OCKHAMG-CDNA  
 INTR 40861 ... 41275 OCKHAMG-CDNA

INIT 40233 ... 40516 OCKHAMG-CDS  
 TERM 40861 ... 41077 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 141  
 - Ceres seq\_id 2047214

(B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 142  
 - Ceres seq\_id 2047215  
 - Location of start within SEQ ID NO 141: at 140 nt.

(B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 143  
 - Ceres seq\_id 2047216  
 - Location of start within SEQ ID NO 141: at 224 nt.

(B) Polypeptide Sequence

321

- Pat. Appln. SEQ ID NO 144
- Ceres seq\_id 2047217
- Location of start within SEQ ID NO 141: at 153 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 144: at 18 aa.

(Ba) Polypeptide Activities: Arabidopsis specific gene, dicot

- 10 specific gene, plant

Maximum Length Sequence:  
Public Genomic DNA:

gi No: 3985958

Predicted Exons:

- 15 INTR 8598 ... 8696 OCKHAMG-CDNA
- INTR 8960 ... 9035 OCKHAMG-CDNA
- INTR 9202 ... 9513 OCKHAMG-CDNA
- INTR 9628 ... 9669 OCKHAMG-CDNA
- 20 INIT 8997 ... 9035 OCKHAMG-CDS
- INTR 9202 ... 9513 OCKHAMG-CDS
- TERM 9628 ... 9822 OCKHAMG-CDS

- 25 (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 145
- Ceres seq\_id 2047438

(B) Polypeptide Sequence

- 30 - Pat. Appln. SEQ ID NO 146
- Ceres seq\_id 2047439
- Location of start within SEQ ID NO 145: at 137 nt.

Maximum Length Sequence:  
Public Genomic DNA:

gi No: 4757403

Predicted Exons:

- 35 INTR 35281 ... 34781 OCKHAMG-CDNA
- 40 SINGLE 35235 ... 35005 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 147
- Ceres seq\_id 2049056

(B) Polypeptide Sequence

- 45 - Pat. Appln. SEQ ID NO 148
- Ceres seq\_id 2049057
- Location of start within SEQ ID NO 147: at 47 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 148: at 22 aa.

322

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 149
- Ceres seq\_id 2049058
- Location of start within SEQ ID NO 147: at 59 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 149: at 18 aa.

Maximum Length Sequence:

Public Genomic DNA:

gi No: 4538972

Predicted Exons:

- 10 INTR 32316 ... 32561 OCKHAMG-CDNA
- INTR 32643 ... 32715 OCKHAMG-CDNA
- INTR 33489 ... 33697 OCKHAMG-CDNA
- INTR 33789 ... 33935 OCKHAMG-CDNA
- INTR 34034 ... 34365 OCKHAMG-CDNA
- 15 INTR 32328 ... 32561 OCKHAMG-CDS
- INTR 32643 ... 32715 OCKHAMG-CDS
- INTR 33489 ... 33697 OCKHAMG-CDS
- INTR 33789 ... 33935 OCKHAMG-CDS
- TERM 34034 ... 34159 OCKHAMG-CDS

- 20 (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 150
- Ceres seq\_id 2050386

(B) Polypeptide Sequence

- 30 - Pat. Appln. SEQ ID NO 151
- Ceres seq\_id 2050387
- Location of start within SEQ ID NO 150: at 1 nt.

(B) Polypeptide Sequence

- 35 - Pat. Appln. SEQ ID NO 152
- Ceres seq\_id 2050388
- Location of start within SEQ ID NO 150: at 13 nt.

(B) Polypeptide Sequence

- 40 - Pat. Appln. SEQ ID NO 153
- Ceres seq\_id 2050389
- Location of start within SEQ ID NO 150: at 151 nt.

(Ba) Polypeptide Activities: Arabidopsis specific gene,

- 45 dicot
- specific gene, plant

Maximum Length Sequence:

Public Genomic DNA:

gi No: 4878038



323

## Predicted Exons:

INTR 22640 ... OCKHANG-CDNA  
 INTR 22412 ... OCKHANG-CDNA  
 INTR 22122 ... OCKHANG-CDNA

5

INTR 22674 ... OCKHANG-CDNA  
 INTR 22412 ... OCKHANG-CDNA  
 INTR 22122 ... OCKHANG-CDNA

10

INIT 22587 ... OCKHANG-CDS  
 INTR 22412 ... OCKHANG-CDS  
 TERM 22122 ... OCKHANG-CDS

gi No: 6143825

## Predicted Exons:

INTR 22639 ... OCKHANG-CDNA  
 INTR 22411 ... OCKHANG-CDNA  
 INTR 22121 ... OCKHANG-CDNA

15

INTR 22673 ... OCKHANG-CDNA  
 INTR 22411 ... OCKHANG-CDNA  
 INTR 22121 ... OCKHANG-CDNA

20

INIT 22586 ... OCKHANG-CDS  
 INTR 22411 ... OCKHANG-CDS  
 TERM 22121 ... OCKHANG-CDS

25

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 154  
 - Ceres seq\_id 2053353

30

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 155  
 - Ceres seq\_id 2053354  
 - Location of start within SEQ ID NO 154: at 88 nt.

35

(Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot

specific gene, plant

specific gene.

40

## Maximum Length Sequence:

Public Genomic DNA:

gi No: 4115352

## Predicted Exons:

INIT 29289 ... OCKHANG-CDS  
 INTR 27681 ... OCKHANG-CDS  
 INTR 27312 ... OCKHANG-CDS  
 INTR 26979 ... OCKHANG-CDS  
 TERM 26654 ... OCKHANG-CDS

45

50

324

(Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 156  
 - Ceres seq\_id 2055693

5

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 157  
 - Ceres seq\_id 2055694  
 - Location of start within SEQ ID NO 156: at 1 nt.

10

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 158  
 - Ceres seq\_id 2055695  
 - Location of start within SEQ ID NO 156: at 46 nt.

15

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 159  
 - Ceres seq\_id 2055696  
 - Location of start within SEQ ID NO 156: at 2 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 159: at 28 aa.

20

(Ba) Polypeptide Activities: Arabidopsis specific gene, dicot

specific gene, plant

specific gene.

25

## Maximum Length Sequence:

Public Genomic DNA:

gi No: 4406790

## Predicted Exons:

INTR 81640 ... 81247 OCKHANG-CDNA  
 SINGLE 81603 ... 81367 OCKHANG-CDS

30

35

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 160  
 - Ceres seq\_id 2056405

40

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 161  
 - Ceres seq\_id 2056406  
 - Location of start within SEQ ID NO 160: at 38 nt.

45

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 162  
 - Ceres seq\_id 2056407  
 - Location of start within SEQ ID NO 160: at 56 nt.

50

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 163

325

- Ceres seq\_id 2056408
- Location of start within SEQ ID NO 160: at 177 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 163: at 14 aa.

5 dicot (Ba) Polypeptide Activities: Arabidopsis specific gene,

specific gene, plant

10

Maximum Length Sequence:  
related to:  
Clone IDs:

233233

15 (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 164
- Ceres seq\_id 2065747

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 165
- Ceres seq\_id 2065748
- Location of start within SEQ ID NO 164: at 114 nt.

25

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 166
- Ceres seq\_id 2065749
- Location of start within SEQ ID NO 164: at 279 nt.

30

(Ba) Polypeptide Activities: Arabidopsis specific gene,

specific gene, plant

326

## SEQ TABLE 1

(2) INFORMATION FOR SEQ ID NO:1:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 601 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (11) MOLECULE TYPE: DNA (genomic)
- (12) FEATURE:
  - (A) NAME/KEY: -
  - (B) LOCATION: 1..601
  - (D) OTHER INFORMATION: SEQ ID NO:1:

10

(D) OTHER INFORMATION: SEQ ID NO:1:

15 gaggaacagc ccaagcgtga ctacatacat aatctcggtt cctctttttt ttgtgtcatt  
ctgttccttc ttgtgttttt gacatttggt ctggaagaaga tgaatttcgtt agaacacgac

15

atggaagatc tgttcacaga gaagaagcgt gtcaagaatc cttctgttcc ttctgtgtgca

20

ctctatgacc cggagagttc. cccgctctgg ttgatttatt tgaagaagc caattccag

25

gtcgttcacg ttctgataag agctcaggtg gtctccacag gtctactcgt cactctaatg

30

gcgtttcac ataatggaga tgaactcgg gaactcgg gaactcgg tgggtgtcct

35

cgctctgagt ttgaagaat cggctctgg taagatttag taaggatata ttactgatt

40

tgaactatc gaagaacgaa caagatcgt tgaactcga tcaattgatt agactgattc

45

aagaaatc tccgaatcc attcttacc tactgtcaa aatctcatra tcttcttgc

50

(2) INFORMATION FOR SEQ ID NO:2:

55

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 114 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear
- (11) MOLECULE TYPE: peptide
- (12) FEATURE:
  - (A) NAME/KEY: peptide
  - (B) LOCATION: 1..114
  - (D) OTHER INFORMATION: / Ceres Seq. ID 1007547

60

(11) SEQUENCE DESCRIPTION: SEQ ID NO:2:

- 1 Asp Glu Gln Ala Thr Arg Asp Tyr Ile His Asn Leu Val Ser Leu Phe
- 5 Phe Phe Val Leu Leu Val Pro Leu Tyr Phe Leu Thr Leu Val Leu Glu
- 10 Lys Met Ser Ser Val Glu Pro Asp Met Glu Asp Leu Phe Gln Lys
- 15 Lys Arg Val Arg Asn Pro Leu Val Pro Leu Gly Ala Leu Met Thr Ala
- 20 Gly Val Leu Thr Ala Gly Leu Ile Ser Phe Arg Arg Gly Asn Ser Gln
- 25 Leu Gly Gln Val Leu Met Arg Ala Arg Val Val Val Gln Gly Ala Thr
- 30 Val Ala Leu Met Val Gly Thr Gly Tyr Tyr Tyr Gly Asp Asn Pro Thr
- 35 Lys Lys

65

(2) INFORMATION FOR SEQ ID NO:3:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 81 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear
- (11) MOLECULE TYPE: peptide
- (12) FEATURE:
  - (A) NAME/KEY: peptide
  - (B) LOCATION: 1..81

327

(D) OTHER INFORMATION: / Ceres Seq. ID 1007548

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Ser Val Glu Pro Asp Met Glu Asp Leu Phe Gln Glu Lys Lys  
 1 5 10 15  
 Arg Val Arg Asn Pro Leu Val Pro Leu Gly Ala Leu Met Thr Ala Gly  
 20 25 30  
 Val Leu Thr Ala Gly Leu Ile Ser Phe Arg Arg Gly Asn Ser Gln Leu  
 35 40 45  
 Gly Gln Val Leu Met Arg Ala Arg Val Val Val Gln Gly Ala Thr Val  
 50 55 60  
 Ala Leu Met Val Gly Thr Gly Tyr Tyr Tyr Gly Asp Asn Pro Trp Lys  
 65 70 75 80  
 Lys

15 (2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 74 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..74

(D) OTHER INFORMATION: / Ceres Seq. ID 1007549

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Glu Asp Leu Phe Gln Glu Lys Lys Arg Val Arg Asn Pro Leu Val  
 1 5 10 15  
 Pro Leu Gly Ala Leu Met Thr Ala Gly Val Leu Thr Ala Gly Leu Ile  
 20 25 30  
 Ser Phe Arg Arg Gly Asn Ser Gln Leu Gly Gln Val Leu Met Arg Ala  
 35 40 45  
 Arg Val Val Val Gln Gly Ala Thr Val Ala Leu Met Val Gly Thr Gly  
 50 55 60  
 Tyr Tyr Tyr Gly Asp Asn Pro Trp Lys Lys  
 65 70

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 970 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..970

(D) OTHER INFORMATION: / Ceres Seq. ID 1007553

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

aacccgctc attttaggg ttacagatc aggcagatc tcacattgc aaacatgacg 60  
 actcagatca gcaagaagc aagttctga gcgagcgtg tctctacgc tgaattgaat 120  
 gaggtctta caagagact agcagaggt ggttacttg ggttgggt taggttact 180  
 cttgagga ctgagattat catcagact acagtactc agaatgtct cggtagaag 240  
 gggagagaa ttagggaatt gactccct gttcagaaga acagaggtc tctgscat cgtcaagc 300  
 agtttgagc ttatgcgca gaaggttaac aacagaggtc tctgscat cgtcaagc 360  
 gacttttac gttaacagct tctgttgt ctgcgttc gtaggttg ctgtgtgtg 420  
 ttgaggttg ttatgagag tggagctaa ggtatgcagg ttatcggag tggaaagct 480  
 cgtgtgcca gacccaagt latgaagtc aagatggtc acatggtc atgtgtcaa 540  
 ccaactaag aatacaga ctctcagtg agactgtt tcttagaca agtgtgttg 600  
 ggaatacag taaagttat cttgtatgg gacctaaag gcatatcag accaagaca 660  
 ccttcyctg atgttgtat cttcttct cctaaagag aagagggcat cctgacact 720  
 gctcaggtg ctgcccggc tgcctcga gcagatgac cactcagc cgtagattac 780  
 cctgcgatg tccagtcgc ctaaagaat ctttgtact agtcagtg agttgttt 840

328

caactgtct tgctctttg acagattata gcttcctga tttcgaatc agacagctt 900  
 cctcgtctc tattgaaaa ttatctcaag ttctcatgt tgaaaaaacc ctttttagt 960  
 tctaaatcg

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 249 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..249

(D) OTHER INFORMATION: / Ceres Seq. ID 1007584

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Thr Thr Gln Ile Ser Lys Lys Arg Lys Phe Val Ala Asp Gly Val  
 1 5 10 15  
 Phe Tyr Ala Glu Leu Asn Glu Val Leu Thr Arg Glu Leu Ala Glu Asp  
 20 25 30  
 Gly Tyr Ser Gly Val Glu Val Arg Val Thr Pro Met Arg Thr Glu Ile  
 35 40 45  
 Ile Ile Arg Ala Thr Arg Thr Gln Asn Val Leu Gly Glu Lys Gly Arg  
 50 55 60  
 Arg Ile Arg Glu Leu Thr Ser Leu Val Gln Lys Arg Phe Lys Phe Pro  
 65 70 75 80  
 Val Asp Ser Val Glu Leu Tyr Ala Glu Lys Val Asn Asn Arg Gly Leu  
 85 90 95  
 Cys Ala Ile Ala Gln Ala Glu Ser Leu Arg Tyr Lys Leu Leu Gly Gly  
 100 105 110  
 Leu Ala Val Arg Arg Ala Cys Tyr Gly Val Leu Arg Phe Val Met Glu  
 115 120 125  
 Ser Gly Ala Lys Gly Cys Glu Val Ile Val Ser Gly Lys Leu Arg Ala  
 130 135 140  
 Ala Arg Ala Lys Ser Met Lys Phe Lys Asp Gly Tyr Met Val Ser Ser  
 145 150 155 160  
 Gly Gln Pro Thr Lys Glu Tyr Ile Asp Ser Ala Val Arg His Val Leu  
 165 170 175  
 Leu Arg Gln Gly Val Leu Gly Ile Lys Val Lys Val Met Leu Asp Trp  
 180 185 190  
 Asp Pro Lys Gly Ile Ser Gly Pro Lys Thr Pro Leu Pro Asp Val Val  
 195 200 205  
 Ile Ile His Ser Pro Lys Glu Glu Glu Ala Ile Tyr Ala Pro Ala Gln  
 210 215 220  
 Val Ala Ala Pro Ala Leu Val Ala Asp Ala Pro Leu Thr Ala Val  
 225 230 235 240  
 Asp Tyr Pro Ala Met Ile Pro Val Ala

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 206 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..206

(D) OTHER INFORMATION: / Ceres Seq. ID 1007585

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Arg Thr Glu Ile Ile Arg Ala Thr Arg Thr Gln Asn Val Leu  
 1 5 10 15  
 Gly Glu Lys Gly Arg Arg Ile Arg Gln Leu Thr Ser Leu Val Gln Lys

329  
20 Arg Phe Lys Phe Pro Val Asp Ser Val Glu Leu Tyr Ala Glu Lys Val 30  
35 Lys Lys Lys Phe Val Cys Ala Ile Ala Glu Ala Glu Ser Leu Arg Tyr 45  
50 Asn Asn Arg Gly Leu Cys Ala Ile Ala Glu Ala Glu Ser Leu Arg Tyr 60  
5 Lys Lys Leu Leu Gly Gly Leu Ala Val Arg Arg Ala Cys Tyr Gly Val Leu 65  
65 Arg Phe Val Met Glu Ser Gly Ala Lys Lys Gly Cys Glu Val Ile Val Ser 80  
10 Gly Lys Leu Arg Ala Ala Arg Ala Lys Ser Met Lys Phe Lys Asp Gly 95  
100 Tyr Met Val Ser Ser Gly Glu Pro Thr Lys Glu Tyr Ile Asp Ser Ala 110  
115 Val Arg His Val Leu Leu Arg Glu Gly Val Leu Gly Ile Lys Val Lys 125  
130 Val Met Leu Asp Trp Asp Pro Lys Gly Ile Ser Gly Pro Lys Thr Pro 140  
145 Leu Pro Asp Val Val Ile Ile His Ser Pro Lys Glu Glu Ala Ile 155  
165 Tyr Ala Pro Ala Glu Val Ala Ala Pro Ala Ala Leu Val Ala Asp Ala 175  
180 Pro Leu Thr Ala Val Asp Tyr Pro Ala Met Ile Pro Val Ala 190  
195  
200  
(2) INFORMATION FOR SEQ ID NO:8:  
(A) LENGTH: 766 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(11) MOLECULE TYPE: DNA (genomic)  
(1x) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..766  
(D) OTHER INFORMATION: / Ceres Seq. ID 1008148  
35 (x1) SEQUENCE DESCRIPTION: SEQ ID NO:8:  
60 aaaaaaac tagtttttg cctctcgaq gttccgtgag aactccgtcg cgaagaagat  
120 cctctaccga tacaatgat atctcagag gtaaccgaa agaatctgc agagactct  
180 tcaagaagag actttctctt gcgaagag attcaatct tgcataagat ccgtcagct  
240 atccaccaga cctcaagatg attaagctta tgcagagtt caatcaagag gagatgta  
300 gagaaactt tgcctgagtg catcaatct ggtttctgac caatcaagag atcgagctc  
360 tgcgaactta tcttaactct catctgagtg ttgtctctgc tactttgaag agtcagcta  
420 atccctgttg tgcatactt gttgcccac ttgtgacgc ctcaagagga cctgcagag  
480 aagggagaga cctccaccag ttgtgtgacg ttgtgtgacg cctgcagag cctgcagag  
540 gttgtgagtg tgcagtgaga aaggtgagag atctcagaa ttacagaca ccttccaa  
600 gaaagtggag tgg-tttgag cgtgtgtcgt gttgtcagag cgcagctgca ccatcgtgt  
660 caggtgtgac ttgaaaaa aatgcttta gttgacaga agacatgga ggaattctta  
720 gcttaaatat ttgttttgt aattcaatc cctgaagctt tcaatactc tatctgagtt  
780 ttgttttgt gttgataca aactcagtt taaagtatg ttcaatcc  
(2) INFORMATION FOR SEQ ID NO:9:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 203 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(11) MOLECULE TYPE: peptide  
(1x) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..203  
(D) OTHER INFORMATION: / Ceres Seq. ID 1008149  
60 (x1) SEQUENCE DESCRIPTION: SEQ ID NO:9:  
1 Lys Asn Pro Ser Phe Leu Leu Ser Arg Ser Ser Val Arg Leu Ser Ala  
5 10 15

330  
20 Arg Ala Arg Ser Pro Thr Asp Thr Met Ile Ile Ser Glu Ala Asn Arg 30  
35 Lys Glu Ile Cys Lys Tyr Leu Phe Lys Glu Gly Val Cys Phe Ala Lys 45  
50 Lys Asp Phe Asn Leu Ala Lys His Pro Leu Ile Asp Val Pro Asn Leu 60  
5 Lys Asp Phe Asn Leu Ala Lys His Pro Leu Ile Asp Val Pro Asn Leu 65  
70 Glu Thr Phe Ala Trp Met His Tyr Tyr Trp Phe Leu Thr Asn Glu Gly 80  
85 Ile Glu Phe Leu Arg Thr Tyr Leu Asn Leu Pro Ser Asp Val Val Pro 95  
100 Ala Thr Lys Lys Ser Ala Lys Pro Gly Gly Arg Pro Phe Gly Gly 110  
115 Pro Pro Gly Asp Arg Ser Arg Gly Pro Arg His Glu Gly Asp Arg 125  
130 Pro Arg Phe Gly Asp Arg Asp Gly Tyr Arg Ala Gly Pro Arg Ala Gly 140  
145 Gly Glu Phe Gly Gly Glu Lys Gly Gly Ala Pro Ala Asp Tyr Glu Pro 150  
155 Ser Phe Glu Gly Ser Gly Arg Gly Phe Gly Arg Gly Ala Gly Gly Tyr 165  
180 Ser Ala Ala Ala Pro Ser Gly Ser Gly Leu Pro 185  
195  
200  
(2) INFORMATION FOR SEQ ID NO:10:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 179 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(11) MOLECULE TYPE: peptide  
(1x) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..179  
(D) OTHER INFORMATION: / Ceres Seq. ID 1008150  
35 (x1) SEQUENCE DESCRIPTION: SEQ ID NO:10:  
60 Met Ile Ile Ser Glu Ala Asn Arg Lys Lys Lys Tyr Leu Phe 15  
1 Lys Glu Gly Val Cys Phe Ala Lys Lys Asp Phe Asn Leu Ala Lys His 25  
30 Pro Leu Ile Asp Val Pro Asn Leu Glu Val Ile Lys Leu Met Glu Ser 35  
40 Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His Tyr 45  
50 Tyr Trp Phe Leu Thr Asn Glu Gly Ile Glu Phe Leu Arg Thr Tyr Leu 55  
60 Asn Leu Pro Ser Asp Val Val Pro Ala Thr Leu Lys Lys Ser Ala Lys 65  
70 Pro Gly Gly Arg Pro Phe Gly Gly Pro Pro Gly Asp Arg Ser Arg Gly 75  
80 Pro Arg His Glu Gly Asp Arg Pro Arg Phe Gly Asp Arg Asp Gly 85  
90 Tyr Arg Ala Gly Pro Arg Ala Gly Gly Glu Phe Gly Gly Glu Lys Gly 95  
100 Gly Ala Pro Ala Asp Tyr Glu Pro Ser Phe Glu Gly Ser Gly Arg Gly 105  
110 Phe Gly Arg Gly Ala Gly Gly Tyr Ser Ala Ala Ala Pro Ser Gly Ser 115  
120 Gly Leu Pro 125  
(2) INFORMATION FOR SEQ ID NO:11:  
(1) SEQUENCE CHARACTERISTICS:

331

(A) LENGTH: 134 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..134

(D) OTHER INFORMATION: / Ceres Seq. ID 1008151

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Gln Ser Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp

1 5 10 15

Met His Tyr Tyr Trp Phe Leu Thr Asn Glu Gly Ile Glu Phe Leu Arg

20 25 30

Thr Tyr Leu Asn Leu Pro Ser Asp Val Val Pro Ala Thr Leu Lys Lys

35 40 45

Ser Ala Lys Pro Gly Gly Arg Pro Phe Gly Gly Pro Pro Gly Asp Arg

50 55 60

Ser Arg Gly Pro Arg His Glu Gly Asp Arg Pro Arg Phe Gly Asp

65 70 75 80

Arg Asp Gly Tyr Arg Ala Gly Pro Arg Ala Gly Gly Glu Phe Gly

85 90 95

Glu Lys Gly Gly Ala Pro Ala Asp Tyr Glu Pro Ser Phe Glu Gly Ser

100 105 110 115

Gly Arg Gly Phe Gly Arg Gly Ala Gly Gly Tyr Ser Ala Ala Pro

120 125

Ser Gly Ser Gly Leu Pro

130

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 995 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..995

(D) OTHER INFORMATION: / Ceres Seq. ID 1008334

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

atattcaaaa gactcaaaaa cctatctca tttatctca acagggcgaa gaacggcgt

60

ggccgccatt aagctctca cctctcgtc tctgtttt ttctcttc catctcgt

120

cacaagccct agattccatg gcttcacca tcaaccacca cagctcgtc ttcttttc

180

tttaacccc ttctcttc tacaatcca aaacctcaaa tccatccgaa ttctcgatc

240

cgcttcgccc ataaacac caatctcca gacggagaag tcaacagtc ggtcatgac

300

acccacggg tccatcgtt cctctctac tctcggct ttagcaatcg ctgtaacag

360

ggtccagct cagaacctt cctctcgaat cagacttca gctccgtaa tgcggacgg

420

gtctcagtc tctctcagta cctccggacc tgtctcttc ggtctctcc gggacgtcc

480

tccggagac ttgaacacg cgtgcaggt tctcgggt gggataaga agtggctaga

540

cattacagt gggatctga tggtaggt tttgtagt tggctcca ataccctg

600

ggaaagacg cctctctc ctattaga tctctgat ccttaactga atctctcag

660

aaacattcat cctctctat cctatcagct tgaatagt ccaatggg tagaattgg aagacttag

720

tcttgataa ctggatcga tggatagg cagactgg tagaattgg aagacttag

780

ctaattgtt ttgtgaacg tcaattgg aaattctgc aactgttc ccaattag

840

ggaaataca gaattgatg ctgtgtga tgcgatga gagtgtgag ctattcgtt

900

ttgtgttc ttgttgac ttgtatcg aataatga ttgtgggtg ttgtttact

960

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 253 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

332

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..253

(D) OTHER INFORMATION: / Ceres Seq. ID 1008335

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Tyr Leu Lys Asp Ser Lys Thr Tyr Arg His Leu Ser Ser Thr Gly Gly

1 5 10 15

Arg Thr Ala Met Ala Ala Ile Thr Ala Leu Thr Leu Arg Ser Pro Val

20 25 30

Phe Leu Leu Pro Pro Ser Ser Val Thr Ser Pro Arg Phe His Gly Phe

35 40 45

Thr Asn Gln Pro Pro Ala Arg Leu Phe Phe Pro Leu Asn Pro Phe

50 55 60

Pro Ser Leu Ser Ile Gln Asn Pro Lys Ser Ile Arg Ile Ser Ala Ser

65 70 75 80

Ala Ser Pro Ile Thr Thr Pro Ile Leu Gln Thr Glu Lys Ser Thr Ala

85 90 95

Arg Ser Ser Thr Leu Thr Gly Ser Thr Arg Ser Leu Ala Thr Leu Ala

100 105 110

Ala Leu Ala Ile Ala Val Thr Arg Val Leu Ala Gln Lys Leu Ser Leu

115 120 125

Ala Ile Gln Thr Ser Ser Pro Val Ile Ala Asp Gly Leu Arg Phe Ser

130 135 140

Leu Ser Thr Ala Gly Pro Val Phe Phe Ala Ser Leu Arg Asp Arg Pro

145 150 155

Pro Gly Tyr Leu Asn Thr Pro Leu Thr Val Val Ala Val Gly Ile Lys

160 165 170

Lys Trp Leu Asp Ile Tyr Ser Gly Val Leu Met Val Arg Val Leu Leu

175 180 185

Ser Trp Phe Pro Asn Ile Pro Trp Glu Arg Gln Pro Leu Ser Ala Ile

190 195 200

Arg Asp Leu Cys Asp Pro Tyr Leu Asn Leu Phe Arg Asn Ile Ile Pro

205 210 215

Pro Ile Phe Asp Thr Leu Asp Val Ser Pro Leu Leu Ala Phe Ala Val

220 225 230

Leu Gly Thr Leu Gly Ser Ile Val His Gly Ser Thr Gly

235 240 245

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 234 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..234

(D) OTHER INFORMATION: / Ceres Seq. ID 1008336

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Ala Ala Ile Thr Ala Leu Thr Leu Arg Ser Pro Val Phe Leu Leu

1 5 10 15

Pro Pro Ser Ser Val Thr Ser Pro Arg Phe His Gly Phe Thr Asn Gln

20 25 30

Pro Pro Ala Arg Leu Phe Phe Pro Leu Asn Pro Phe Pro Ser Leu

35 40 45

Ser Ile Gln Asn Pro Lys Ser Ile Arg Ile Ser Ala Ser Ala Ser Pro

50 55 60

Ile Thr Thr Pro Ile Leu Gln Thr Glu Lys Ser Thr Ala Arg Ser Ser

65 70 75 80

Thr Leu Thr Gly Ser Thr Arg Ser Leu Ala Thr Leu Ala Ala Leu Ala

90 95 100 105

333  
05 90 95  
11e Ala Val Thr Arg Val Leu Ala Gln Lys Leu Ser Leu Ala Ile Gln  
100 105 110  
Thr Ser Ser Pro Val Ile Ala Asp Gly Leu Arg Phe Ser Leu Ser Thr  
5 115 120 125  
Ala Gly Pro Val Phe Ala Ser Leu Arg Asp Arg Pro Pro Gly Tyr  
130 135 140  
Leu Asn Thr Pro Leu Thr Val Val Ala Val Gly Ile Lys Lys Trp Leu  
145 150 155 160 165  
Asp Ile Tyr Ser Gly Val Leu Met Val Arg Val Leu Leu Ser Trp Phe  
170 175  
Pro Asn Ile Pro Trp Gln Arg Gln Pro Leu Ser Ala Ile Arg Asp Leu  
180 185 190 195  
Cys Asp Pro Tyr Leu Asn Leu Phe Arg Asn Ile Ile Pro Pro Ile Phe  
200 205  
Asp Thr Leu Asp Val Ser Pro Leu Leu Ala Phe Ala Val Leu Gly Thr  
210 215 220  
Leu Gly Ser Ile Val His Gly Ser Thr Gly  
225 230  
(1) INFORMATION FOR SEQ ID NO: 5:  
(2) INFORMATION FOR SEQ ID NO: 5:  
(A) LENGTH: 523 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(E) MOLECULE TYPE: DNA (genomic)  
(F) FEATURE:  
(G) NAME/KEY: -  
(H) LOCATION: 1..520  
(I) OTHER INFORMATION: / Ceres Seq. ID 1008701  
30  
(X1) SEQUENCE DESCRIPTION: SEQ ID NO:15:  
acacataacc cctctggaga gagctccaac actcgagag aaaaagaaga aggaagatcc  
cgaagaatg caaagcgat tgaagttca gctcttccc gaggacgagc tggcgagct  
cgaagacat ccgctgcac taaggaaac tttcttcct cgcgcgcga tgcagatgt  
tatgaacgt cgaagtgga gaagatcat tcttgagga lggctagttg ccttgctta  
gcgcctacg ttatattaca ggtacacat ccgcgcgag acccctctgc ctacgcgat  
atgcacatg actgaagat gctgcttgg gttcccgatg gtctgttga gttgaagac  
aaacaagag actgaatct gcgttgcat aataagctt tcttggtta ttgaagac  
taaatgtt taccgattt gttccacag ttgtaacag atttctact ccaatctct  
tctcttctt cgaataaaa agtaataac ttgttggtc  
40  
(2) INFORMATION FOR SEQ ID NO:16:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 124 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(E) MOLECULE TYPE: peptide  
(F) FEATURE:  
(G) NAME/KEY: peptide  
(H) LOCATION: 1..124  
(I) OTHER INFORMATION: / Ceres Seq. ID 1008702  
50  
(X1) SEQUENCE DESCRIPTION: SEQ ID NO:16:  
Thr His Asn Pro Leu Gly Lys Ser Leu Asn Thr Cys Arg Gln Lys Gln  
1 5 10 15  
Glu Gly Arg Ser Arg Lys Met Ala Thr Ala Ile Val Arg Ser Ala Leu  
20 25 30  
Ser Arg Ala Ala Thr Arg Ala Ala Pro Lys Thr Ser Val Ala Pro Lys  
35 40 45  
Arg Asn Phe Ser Ser Ala Gly His Asp Asp Ala Tyr Gln Ala Ala  
50 55 60  
Lys Trp Gln Lys Ile Thr Tyr Leu Gly Ile Ala Ser Cys Thr Ala Leu  
65 70 75 80

334  
Ala Val Tyr Val Leu Ser Lys Gly His His Pro Gly Gln Asp Pro Pro  
85 90 95  
Ala Tyr Pro His Met His Ile Arg Asn Lys Gln Phe Pro Trp Gly Pro  
100 105 110  
Asp Gly Leu Phe Gln Val Lys His Asn Lys Gln His  
115 120  
(2) INFORMATION FOR SEQ ID NO:17:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 102 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(E) MOLECULE TYPE: peptide  
(F) FEATURE:  
(G) NAME/KEY: peptide  
(H) LOCATION: 1..102  
(I) OTHER INFORMATION: / Ceres Seq. ID 1008703  
15  
(X1) SEQUENCE DESCRIPTION: SEQ ID NO:17:  
Met Ala Thr Ala Ile Val Arg Ser Ala Leu Ser Arg Ala Ala Thr Arg  
1 5 10 15  
Ala Ala Pro Lys Thr Ser Val Ala Pro Lys Arg Asn Phe Ser Ser  
20 25  
Ala Gly His Asp Asp Ala Tyr Gln Ala Ala Lys Trp Gln Lys Ile Thr  
30 35  
Tyr Leu Gly Ile Ala Ser Cys Thr Ala Leu Ala Val Tyr Val Leu Ser  
40 45  
Lys Gly His His Pro Gly Gln Asp Pro Pro Ala Tyr Pro His Met His  
50 55 60  
Ile Arg Asn Lys Gln Phe Pro Trp Gly Pro Asp Gly Leu Phe Gln Val  
65 70 75 80 85 90 95  
Lys His Asn Lys Gln His  
100  
(2) INFORMATION FOR SEQ ID NO:18:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 68 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(E) MOLECULE TYPE: peptide  
(F) FEATURE:  
(G) NAME/KEY: peptide  
(H) LOCATION: 1..68  
(I) OTHER INFORMATION: / Ceres Seq. ID 1008704  
45  
(X1) SEQUENCE DESCRIPTION: SEQ ID NO:18:  
His Ile Thr Leu Leu Gln Arg Val Ser Thr Leu Ala Gln Lys Lys Lys  
1 5 10 15  
Lys Gln Asp Pro Gly Lys Trp Gln Arg Arg Leu Tyr Val Gln Leu Phe  
20 25 30  
Pro Gln Gln Arg Leu Ala Gln Leu Arg Arg His Pro Ser Leu Leu Ser  
35 40 45  
Gln Thr Phe Pro Leu Pro Pro Ala Met Thr Met Leu Met Lys Leu Arg  
50 55 60  
Ser Gly Arg Arg  
65  
(2) INFORMATION FOR SEQ ID NO:19:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 662 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(E) MOLECULE TYPE: DNA (genomic)  
(F) FEATURE:

335

(A) NAME/KEY: -  
 (B) LOCATION: 1..662  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1009063  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

5 catctttttt tctctctct cctccctaa gcaaaactaa acaagctatg gctggtatgc 60  
 tctccggagt tgaatgta agagggggc gcttccacg tggctctct ccgattgaat 120  
 cctgaacac agttctg gggctggc gggaacat ctagacacg cgaacatct 180  
 tctctctta taccctaat catgagggc accagggcca tctctcttc tggagagaa 240  
 ggttagaa taactttat gtagaaga acgatgaga actttacgga gcagccaag 300  
 agtcaasga gagctctaac aagcgtctga gaattccacc acgtcaagt tcaagcaaa 360  
 tggtaagac aaaggaata aattggaga aggaaggggt aaactctcg ggaatttacc 420  
 gaccggggt gtcgggttaa agagagggc aggaaggttg atggaatggt tcaagggcg 480  
 agtaggaa caaagatt ggtcatatg tctagacgg tcaagaagg gtagacctt 540  
 gtagaccta caatggccc ataatgttca ctccatagc ttatggcctt ggtagaac 600  
 15 taatgtttat tgcctattt gtagaactga tatttggat taaatgttat attttgatg 660

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 130 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

(A) NAME/KEY: peptide  
 (B) LOCATION: 1..130  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1009064

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

30 Met Ala Gly Met Leu Pro Gly Val Glu Cys Ala Arg Arg Arg Arg Phe 15  
 1 His Gly Gly Ala Pro Ile Glu Ser Ser Asn Thr Ala Ser Val Ala 30  
 Ala Ala Ala Gly His Val Thr Arg Arg Pro Ser Phe Ser Leu Tyr 25  
 35 Thr Thr Asn His Glu Ser His Gln Ala His Val Ser Phe Ser Glu Arg 45  
 50 Ser Val Arg Asn Lys Ser Tyr Gly Glu Asp Asn Asp Glu Lys Leu Asp 60  
 65 Gly Ala Ala Lys Glu Ala Lys Gln Arg Leu Asn Lys Arg Leu Arg Ile 75  
 85 Pro Pro Arg Thr Ser Ser Gly Lys Met Val Lys Thr Lys Gly Ile Asn 90  
 100 Trp Ser Lys Glu Arg Val Asn Leu Ser Gly Thr Tyr Arg Pro Arg Trp 105  
 45 Ser Gly 115 120 125

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 127 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

(A) NAME/KEY: peptide  
 (B) LOCATION: 1..127  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1009065

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

50 Met Leu Pro Gly Val Glu Cys Ala Arg Arg Arg Arg Phe His Gly Gly 15  
 1 Ala Pro Pro Ile Glu Ser Ser Asn Thr Ala Ser Val Ala Ala Ala 30

336

Gly His Val Trp Thr Arg Arg Pro Ser Phe Ser Leu Tyr Thr Thr Asn 45  
 50 His Glu Ser His Gln Ala His Val Ser Phe Ser Glu Arg Ser Val Arg 55  
 60 Asn Lys Ser Tyr Gly Glu Asp Asn Asp Glu Lys Leu Asp Gly Ala Ala 65  
 70 Lys Glu Ala Lys Gln Arg Leu Asn Lys Arg Leu Arg Ile Pro Pro Arg 75  
 85 Thr Ser Ser Gly Lys Met Val Lys Thr Lys Gly Ile Asn Trp Ser Lys 90  
 100 Cys Arg Val Asn Leu Ser Gly Thr Tyr Arg Pro Arg Trp Ser Gly 105  
 115 120 125

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 504 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:

(A) NAME/KEY: -  
 (B) LOCATION: 1..504  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1009345

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

25 aagaacaaa ggtgtcaaga agaagatcca acatattaat tgactaaaa tgaatgcag 60  
 gatataatc ttgtgtgta ttgtgtgcg ttgtgtggt gagcaatcag agctcgcga 120  
 agtgcgaag tatttaaac cagcgtctc tgacgtgtc cgtgtccta atccctcag 180  
 ggaatgcat ccacacatt ccacacaaa acctgcgtc cctgttcaca attaagtcg 240  
 tgggtgatt agaattacc ggtgcagcg agatgcctg gtttaacct acatgcagac 300  
 gactccttt tatattgac ctctcttcac atgagacttg caatgattt atatttcac 360  
 atatacac attacgtgt atagtataa tccagtgtg taatatata atctattga 420  
 ttcaatcca atacacaga tctgttaact gaataattt tgcctcattt ttgtgctat 480  
 aataagga tgaatttat tatt

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 76 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

(A) NAME/KEY: peptide  
 (B) LOCATION: 1..76  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1009346

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

45 Met Asn Ala Trp Met Ile Ile Leu Leu Val Ile Cys Val Ala Val Val 10  
 1 Val Glu Gln Ser Glu Ala Arg Lys Gly Arg Lys Tyr Leu Asn Pro Gly 15  
 20 Val Leu Asp Arg Cys Arg Gly Pro Asn Pro Pro Ala Gly Cys His Pro 25

50 His Asn Ser His Lys Pro Arg Val Pro Val His Asn Tyr Ser Arg 45  
 55 Gly Cys Ser Arg Ile Thr Arg Cys Arg Arg Asp Ala 60  
 65 70 75

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 72 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide

337

## (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..72

(D) OTHER INFORMATION: / Ceres Seq. ID 1009347

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Met Ile Ile Leu Leu Val Ile Cys Val Ala Val Val Val Glu Gln Ser  
1 5 10 15  
Glu Ala Arg Lys Gly Arg Lys Tyr Leu Asn Pro Gly Val Leu Asp Arg  
20 25 30  
Cys Arg Gly Pro Asn Pro Pro Ala Gly Cys His Pro His Asn Ser His  
35 40 45  
His Lys Pro Arg Val Pro Val His Asn Tyr Ser Arg Gly Cys Ser Arg  
50 55 60  
Ile Thr Arg Cys Arg Arg Asp Ala

(2) INFORMATION FOR SEQ ID NO:25:  
(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 967 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..967

(D) OTHER INFORMATION: / Ceres Seq. ID 1010140

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

attctatc tctccacc atctccatc ctctcttc aatgcctcc ttatctctg  
120  
ttctcttc tctccacc atctccatc ctctcttc aatgcctcc ttatctctg  
180  
ctctccacc atctccatc ctctcttc aatgcctcc ttatctctg  
240  
attgcctcc aatgcctcc aatgcctcc aatgcctcc aatgcctcc  
300  
ggctgcgacc cgaatccac ggaatcttc ctgtgtatc gacgggaaa ctactctc  
360  
ccgtgcgacc cgaatccac ggaatcttc ctgtgtatc gacgggaaa ctactctc  
420  
gtgtgcgacc cgaatccac ggaatcttc ctgtgtatc gacgggaaa ctactctc  
480  
acgggcgacc aatgcctcc aatgcctcc aatgcctcc aatgcctcc  
540  
aactgcgacc aatgcctcc aatgcctcc aatgcctcc aatgcctcc  
600  
agctgcgacc aatgcctcc aatgcctcc aatgcctcc aatgcctcc  
660  
agctgcgacc aatgcctcc aatgcctcc aatgcctcc aatgcctcc  
720  
attgcctcc aatgcctcc aatgcctcc aatgcctcc aatgcctcc  
780  
ctcgaagatc ggtgcgacc ggtgcgacc ggtgcgacc ggtgcgacc  
840  
gaaagaagc cgtgcgacc atgcgaagc ggtgcgacc ggtgcgacc  
900  
ttgcgttcac aatctcttc cctgttcac aatctcttc ggtgcgacc  
960

(2) INFORMATION FOR SEQ ID NO:26:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 289 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..289

(D) OTHER INFORMATION: / Ceres Seq. ID 1010141

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Phe Ile Ser Ser Pro Thr Ile Leu Asn Pro Leu Ser Pro Met Ala Ser  
1 5 10 15  
Leu Ser Phe Val Ser Ser His Leu Thr Leu Arg Thr Pro Ser Ile  
20 25 30  
Ala Leu Arg Ser Thr Gly Ser Ser Pro Arg Thr Ser Val Ser Ser  
35 40 45

338

## (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..72

(D) OTHER INFORMATION: / Ceres Seq. ID 1009347

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Met Ile Ile Leu Leu Val Ile Cys Val Ala Val Val Val Glu Gln Ser  
1 5 10 15  
Glu Ala Arg Lys Gly Arg Lys Tyr Leu Asn Pro Gly Val Leu Asp Arg  
20 25 30  
Cys Arg Gly Pro Asn Pro Pro Ala Gly Cys His Pro His Asn Ser His  
35 40 45  
His Lys Pro Arg Val Pro Val His Asn Tyr Ser Arg Gly Cys Ser Arg  
50 55 60  
Ile Thr Arg Cys Arg Arg Asp Ala

(2) INFORMATION FOR SEQ ID NO:25:  
(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 967 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..967

(D) OTHER INFORMATION: / Ceres Seq. ID 1010140

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

attctatc tctccacc atctccatc ctctcttc aatgcctcc ttatctctg  
120  
ttctcttc tctccacc atctccatc ctctcttc aatgcctcc ttatctctg  
180  
ctctccacc atctccatc ctctcttc aatgcctcc ttatctctg  
240  
attgcctcc aatgcctcc aatgcctcc aatgcctcc aatgcctcc  
300  
ggctgcgacc cgaatccac ggaatcttc ctgtgtatc gacgggaaa ctactctc  
360  
ccgtgcgacc cgaatccac ggaatcttc ctgtgtatc gacgggaaa ctactctc  
420  
gtgtgcgacc cgaatccac ggaatcttc ctgtgtatc gacgggaaa ctactctc  
480  
acgggcgacc aatgcctcc aatgcctcc aatgcctcc aatgcctcc  
540  
aactgcgacc aatgcctcc aatgcctcc aatgcctcc aatgcctcc  
600  
agctgcgacc aatgcctcc aatgcctcc aatgcctcc aatgcctcc  
660  
agctgcgacc aatgcctcc aatgcctcc aatgcctcc aatgcctcc  
720  
attgcctcc aatgcctcc aatgcctcc aatgcctcc aatgcctcc  
780  
ctcgaagatc ggtgcgacc ggtgcgacc ggtgcgacc ggtgcgacc  
840  
gaaagaagc cgtgcgacc atgcgaagc ggtgcgacc ggtgcgacc  
900  
ttgcgttcac aatctcttc cctgttcac aatctcttc ggtgcgacc  
960

(2) INFORMATION FOR SEQ ID NO:26:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 276 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..276

(D) OTHER INFORMATION: / Ceres Seq. ID 1010142

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Phe Ile Ser Ser Pro Thr Ile Leu Asn Pro Leu Ser Pro Met Ala Ser  
1 5 10 15  
Leu Ser Phe Val Ser Ser His Leu Thr Leu Arg Thr Pro Ser Ile  
20 25 30  
Ala Leu Arg Ser Thr Gly Ser Ser Pro Arg Thr Ser Val Ser Ser  
35 40 45  
Ala Leu Arg Ser Thr Gly Ser Ser Pro Arg Thr Ser Val Ser Ser  
50 55 60  
Ala Leu Arg Ser Thr Gly Ser Ser Pro Arg Thr Ser Val Ser Ser  
65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 900 905 910 915 920 925 930 935 940 945 950 955 960 965 970 975 980 985 990 995 1000



339  
Ala Leu Leu Arg Glu Lys Met Val Glu Ala Val Ala Asp Lys Phe Ile  
145 150 155  
Val Val Ala Asp Asp Thr Lys Leu Val Thr Gly Leu Gly Ser Gly  
165 170 175  
Leu Ala Met Pro Val Glu Val Val Gln Phe Cys Trp Asn Phe Asn Leu  
180 185 190  
Ile Arg Leu Gln Glu Leu Phe Lys Glu Phe Gly Cys Glu Ser Lys Leu  
195 200  
Arg Val Asp Gly Asp Gly Lys Pro Tyr Val Thr Asp Asn Ser Asn Tyr  
210 215 220  
Ile Ile Asp Leu Tyr Phe Lys Thr Pro Leu Lys Asp Gly Phe Ala Ala  
225 230 235  
Ala Lys Glu Ile Gly Lys Phe Gln Gly Val Val Glu His Gly Leu Phe  
240 245 250 255  
Leu Gly Met Ala Thr Ser Val Ile Ile Ala Gly Lys Asn Gly Val Glu  
260 265 270  
Val Met Thr Lys  
275

## (2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 213 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) FEATURE:  
(i) MOLECULE TYPE: peptide  
(ii) NAME/KEY: peptide  
(B) LOCATION: 1..213  
(D) OTHER INFORMATION: / Ceres Seq. ID 1010143

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Met Val Leu Gly Leu Gly Thr Gly Ser Thr Ala Ala Phe Ala Val Asp  
1 5 10 15  
Gln Ile Gly Lys Leu Leu Ser Ser Gly Glu Leu Tyr Asp Ile Val Gly  
20 25 30  
Ile Pro Thr Ser Lys Arg Thr Glu Glu Gln Ala Arg Ser Leu Gly Ile  
35 40 45  
Pro Leu Val Gly Leu Asp Thr His Pro Arg Ile Asp Leu Ala Ile Asp  
50 55 60  
Gly Ala Asp Glu Val Asp Pro Asn Ieu Asp Leu Val Lys Gly Arg Gly  
65 70 75 80  
Gly Ala Leu Leu Arg Glu Lys Met Val Glu Ala Val Ala Asp Lys Phe  
85 90 95  
Ile Val Val Ala Asp Asp Thr Lys Leu Val Thr Gly Leu Gly Gly Ser  
100 105 110  
Gly Leu Ala Met Pro Val Glu Val Val Gln Phe Cys Trp Asn Phe Asn  
115 120  
Leu Ile Arg Leu Gln Glu Leu Phe Lys Glu Phe Gly Cys Glu Ser Lys  
125 130 135  
Leu Arg Val Asp Gly Asp Gly Lys Pro Tyr Val Thr Asp Asn Ser Asn  
140 145 150 155  
Tyr Ile Ile Asp Leu Tyr Phe Lys Thr Pro Leu Lys Asp Gly Phe Ala  
160 165 170  
Ala Ala Lys Glu Ile Gly Lys Phe Gln Gly Val Val Glu His Gly Leu  
175 180 185 190  
Phe Leu Gly Met Ala Thr Ser Val Ile Ile Ala Gly Lys Asn Gly Val  
195 200 205 210  
Glu Val Met Thr Lys

## (2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 623 base pairs  
(B) TYPE: nucleic acid

340

- (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

5  
(A) NAME/KEY: -  
(B) LOCATION: 1..623  
(D) OTHER INFORMATION: / Ceres Seq. ID 1010217  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:  
Gaaacccctt aagctaaagc aattcaaaag caattcaaaa accctagcct ctctctctt  
10 actctcaca acaaaatttt tgaatggcg attcttaagg ctccattgt tttctcttg  
atgggatta tctccgcgt tgcctggca cagtctgagc cactagcacc agctccact  
120 ttggatcta gtgcatttc agctcttc gtgtcagctg tgcagtcac accgctgag  
atgggaagca ggggtttgag g-atagagag aagaggaaga a-cggaaatt tgaagaag  
240 atagctatg cgcgcgtaa agctacatc agatagagc cagagaatcaa agagcgtttt  
360 gccaagaaa cagattccag agaatgatc ggtggagcg tggagattca ttggcgtatc  
420 ggtgtgtc cagattcttg atattccg gttaaagaa caggtagta gtaatacgt  
480 taatagaga ttaaggataa ttaaaactaa aatgttctc gatttaaca gggctaatg  
540 tattotaga agattctgt ttttgaatt aattccaca attgtatctc aaattatata  
600 gttacttaa tttgaacta atg

## (2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 118 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) FEATURE:  
(i) MOLECULE TYPE: peptide  
(ii) NAME/KEY: peptide  
(B) LOCATION: 1..118  
(D) OTHER INFORMATION: / Ceres Seq. ID 1010213

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Met Ala Ile Ser Lys Ala Ser Ile Val Val Leu Met Met Val Ile Ile  
1 5 10 15  
Ser Val Val Ala Ser Ala Gln Ser Glu Ala Pro Ala Pro Ser Pro Thr  
20 25 30  
Ser Gly Ser Ser Ala Ile Ser Ala Ser Phe Val Ser Ala Val Gln Leu  
35 40  
Ser Pro Ala Glu Arg Glu Ala Arg Val Leu Arg Tyr Arg Lys Arg  
50 55 60  
Lys Asn Arg Lys Phe Glu Lys Thr Ile Arg Tyr Ala Ser Arg Lys Ala  
65 70 75 80  
Tyr Ala Glu Met Arg Pro Arg Ile Lys Gly Arg Phe Ala Lys Arg Thr  
85 90 95  
Asp Ser Arg Glu Asn Asp Gly Gly Asp Val Gly Val Tyr Cys Gly Phe  
100 105 110  
Gly Val Val Pro Ser Phe

## (2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 107 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) FEATURE:  
(i) MOLECULE TYPE: peptide  
(ii) NAME/KEY: peptide  
(B) LOCATION: 1..107  
(D) OTHER INFORMATION: / Ceres Seq. ID 1010219

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Met Met Val Ile Ile Ser Val Val Ala Ser Ala Gln Ser Glu Ala Pro  
1 5 10 15  
Ala Pro Ser Pro Thr Ser Gly Ser Ser Ala Ile Ser Ala Ser Phe Val

20 341  
Ser Ala Val Gln Leu Ser Pro Ala Glu Arg Glu Ala Arg Val Leu Arg  
35 40 45  
Tyr Arg Glu Lys Arg Lys Asn Arg Lys Phe Glu Lys Thr Ile Arg Tyr  
50 55 60  
Ala Ser Arg Lys Ala Tyr Ala Glu Met Arg Pro Arg Ile Lys Gly Arg  
65 70 75 80  
Phe Ala Lys Arg Thr Asp Ser Arg Glu Asn Asp Gly Gly Asp Val Gly  
85 90 95  
Val Tyr Cys Gly Phe Gly Val Val Pro Ser Phe  
100 105  
(2) INFORMATION FOR SEQ ID NO:32:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 106 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(1x) MOLECULE TYPE: peptide  
(1x) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..106  
(D) OTHER INFORMATION: / Ceres Seq. ID 1010220  
(x1) SEQUENCE DESCRIPTION: SEQ ID NO:32:  
Met Val Ile Ile Ser Val Val Ala Ser Ala Gln Ser Glu Ala Pro Ala  
1 5 10 15  
Pro Ser Pro Thr Ser Gly Ser Ser Ala Ile Ser Ala Ser Phe Val Ser  
20 25 30  
Ala Val Gln Leu Ser Pro Ala Glu Arg Glu Ala Arg Val Leu Arg Tyr  
35 40 45  
Arg Glu Lys Arg Lys Asn Arg Lys Phe Glu Lys Thr Ile Arg Tyr Ala  
50 55 60  
Ser Arg Lys Ala Tyr Ala Glu Met Arg Pro Arg Ile Lys Gly Arg Phe  
65 70 75 80  
Ala Lys Arg Thr Asp Ser Arg Glu Asn Asp Gly Gly Asp Val Gly Val  
85 90 95  
Tyr Cys Gly Phe Gly Val Val Pro Ser Phe  
100 105  
(2) INFORMATION FOR SEQ ID NO:33:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 568 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(1x) MOLECULE TYPE: DNA (genomic)  
(1x) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..568  
(D) OTHER INFORMATION: / Ceres Seq. ID 1010302  
(x1) SEQUENCE DESCRIPTION: SEQ ID NO:33:  
aaagaaatc attcaaaagc atacaataa a-ttgagttt -tttttaatt agaaacaat  
60  
gggtttgagc ggtgtcttc atgtgaggt tgaag-taag tctccgagtc aaagtctcg  
120  
gttagccctc gggagagcga tcaatctctt cccaagagc ttccctaaag attcaaac  
180  
cttcaatct cttagcgagc acggaacgc tctcgcttc atcgctcca ttaactatg  
240  
aagagatct ccaatgttga agatctcgc ggaagagatc gaagcattga a-ttggagaa  
300  
caaaagcatg tcaatacaga tcaattgcgc ggaattgtg gaattactca aaacttcaa  
360  
aggaacacac acgttatcct ctaagaacgg ttggaacctt ctgaattgt ctggttgaag  
420  
ttgaagagac gcccaatgaa ttgattgacc aaagctatc aaggaattg ctgcaagaa  
480  
cttaaaagc ataatgagc attctcttaa gcaactatg gcttaact actgaacctt agaacctta  
540  
aactaataa gaagttcga tgcctct  
(2) INFORMATION FOR SEQ ID NO:34:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 155 amino acids

342  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(1x) MOLECULE TYPE: peptide  
(1x) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..155  
(D) OTHER INFORMATION: / Ceres Seq. ID 1010303  
(x1) SEQUENCE DESCRIPTION: SEQ ID NO:34:  
Met Gly Leu Ser Gly Val Leu His Val  
1 5 10  
Ala Glu Lys Phe Tyr Val Ala Leu Gly Asp Gly Ile Asn Leu Phe Pro  
15 20 25 30  
Lys Ala Phe Pro Asn Asp Tyr Lys Thr Ile Gln Val Leu Ala Gly Asp  
35 40 45  
Gly Asn Ala Pro Gly Ser Ile Arg Leu Ile Thr Tyr Gly Glu Gly Ser  
50 55 60  
Pro Leu Val Lys Ile Ser Ala Glu Arg Ile Glu Ala Val Asp Leu Glu  
65 70 75 80  
Asn Lys Ser Met Ser Tyr Ser Ile Ile Gly Gly Glu Met Leu Glu Tyr  
85 90 95  
Tyr Lys Thr Phe Lys Gly Thr Ile Thr Val Ile Pro Lys Asn Gly Gly  
100 105 110  
Ser Leu Leu Lys Thr Ser Gly Glu Phe Glu Lys Thr Ala His Glu Ile  
115 120 125  
Asp Asp Pro His Val Ile Lys Asp Phe Ala Val Lys Asn Phe Lys Glu  
130 135 140  
Ile Asp Glu Tyr Leu Leu Lys Gln Thr Ser Ala  
145 150 155  
(2) INFORMATION FOR SEQ ID NO:35:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 531 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(1x) MOLECULE TYPE: DNA (genomic)  
(1x) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..531  
(D) OTHER INFORMATION: / Ceres Seq. ID 1010815  
(x1) SEQUENCE DESCRIPTION: SEQ ID NO:35:  
ggcgagccgc aaccggaaga agaactctc ctggaacgc ttgga-cctga attcccgag  
60  
cggtgaaata tggttgcgga gaagaagcg aagaagctcc atgaagatc caacgacga  
120  
ttagctcttg tggtagagc tgaagaagtc acctcttgat aagaactgt tctcaatcc  
180  
cttcgagct ccaagagtaa gctgattcg atatcagca atgcacacc gttagaagc  
240  
tcaagattg aataatgac gatctctgt aaagtctgag ttcatctgca caatggcac  
300  
aacttgatt tgggaactgc ttggtgtaa taacttcgag ttctctgct caagctatc  
360  
gactctggcg attcgagat catcaagaa ctctctcgag atcagatcat ctgatttga  
420  
tgattctgc atttgcttc ctatgcttga ttatgcttg tgccttcaa agactaatg  
480  
ttaatttct ttggttccc gactctctc taatgttat atcaattac t  
(2) INFORMATION FOR SEQ ID NO:36:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 112 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(1x) MOLECULE TYPE: peptide  
(1x) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..112  
(D) OTHER INFORMATION: / Ceres Seq. ID 1010816  
(x1) SEQUENCE DESCRIPTION: SEQ ID NO:36:

343

Met Val Ala Glu Lys Lys Ala Lys Lys Ser His Glu Gly Ile Asn Ser  
1 5 10 15  
Arg Leu Ala Leu Val Met Lys Ser Gly Lys Tyr Thr Leu Gly Tyr Lys  
20 25 30  
Ser Val Leu Lys Ser Leu Arg Ser Lys Gly Lys Leu Ile Leu Ile  
35 40 45  
Ser Ser Asn Cys Pro Leu Arg Arg Ser Glu Ile Glu Tyr Tyr Ala  
50 55 60  
Met Leu Ala Lys Val Gly Val His Arg Tyr Asn Gly Asn Asn Val Asp  
65 70 75 80  
Leu Gly Thr Ala Cys Gly Lys Tyr Phe Arg Val Ser Cys Leu Ser Ile  
85 90  
Val Asp Pro Gly Asp Ser Asp Ile Ile Lys Thr Leu Pro Gly Asp Gln  
100 105 110

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 91 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(E) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..91

(D) OTHER INFORMATION: / Ceres Seq. ID 1010817

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Met Lys Ser Gly Lys Tyr Thr Leu Gly Tyr Lys Ser Val Leu Lys Ser  
1 5 10 15  
Leu Arg Ser Ser Lys Gly Lys Leu Ile Leu Ser Ser Asn Cys Pro  
20 25 30  
Pro Leu Arg Arg Ser Glu Ile Glu Tyr Tyr Ala Met Leu Ala Lys Val  
35 40 45  
Gly Val His Arg Tyr Asn Gly Asn Asn Val Asp Leu Gly Thr Ala Cys  
50 55 60  
Gly Lys Tyr Phe Arg Val Ser Cys Leu Ser Ile Val Asp Pro Gly Asp  
65 70 75 80  
Ser Asp Ile Ile Lys Thr Leu Pro Gly Asp Gln  
85 90

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 73 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(E) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..73

(D) OTHER INFORMATION: / Ceres Seq. ID 1010818

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Met Arg Cys Leu Leu Lys Leu Glu Phe Ile Ala Thr Met Ala Thr Met  
1 5 10 15  
Leu Ile Thr Val Leu Leu Val Val Asn Thr Ser Glu Phe Leu Ala Ser  
20 25 30  
Ala Ser Leu Ile Leu Val Ile Pro Thr Ser Ser Arg His Phe Leu Glu  
35 40 45  
Ile Ser Asp Ser Asp Phe Asp Asp Phe Ala Met Leu Phe Ser Met Leu  
50 55 60  
Asp Tyr Ala Leu Cys Ser Ser Lys Thr  
65 70

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

344

(A) LENGTH: 494 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(E) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..494

(D) OTHER INFORMATION: / Ceres Seq. ID 1011437

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

acaatcgga gaaagagaa gacgaagatg tcttggtat gactggaagc gatgtgctt  
ctcggaatc tgggtggat gctctgac atggcaatt ctcagtaata caccacaaa  
gcttaccatg gccctcttaa gacatggc cagatgaat gggatgttc tatggaaga  
cgcacaaga aag-cgtcga gaaatgca gctcttct catgatcgc ttatctctt  
ttgtctct cagggtcta agtgaccac ttggtgac aaataaagtg cattccaga  
gaagagaag ctgggggac tagtaactt atccattc gatttctt ggaatatta  
aagcttcag aaacagacc tcaatacat ttgtttatc aataattc cattctgat  
atttgatg cttttaagt gcatgataa ctgaactact ccttgctgt gtgagtgat  
aaatgaatat ctg.

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 94 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(E) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..94

(D) OTHER INFORMATION: / Ceres Seq. ID 1011438

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

G.c Ser Glu Glu Arg Arg Arg Arg Cys Arg Tyr Tyr Gly Tyr Lys  
1 5 10 15  
Arg Cys Cys Leu Ser Glu Ser Ser Val Gly Cys Ser Val Ser Trp Ala  
20 25 30  
Ile Leu Ser Thr Thr Ser Thr Lys Leu Ile Met Ala Val Leu Ser Thr  
35 40 45  
Ser Ala Thr Met Asn Gly Met Phe Leu Trp Lys Asp Ala Thr Arg Lys  
50 55 60  
Ser Ser Arg Lys Leu Gln Leu Leu Pro His Asp Ser Leu Tyr Leu Phe  
65 70 75 80  
Cys Val Pro Gln Gly Leu Lys Val Thr Thr Cys Gly Asp Lys  
85 90

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 74 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(E) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..74

(D) OTHER INFORMATION: / Ceres Seq. ID 1011439

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Thr Ile Gly Arg Lys Glu Lys Thr Lys Met Ser Leu Val Trp Leu Glu  
1 5 10 15  
Ala Met Leu Pro Leu Gly Ile Ile Gly Gly Met Leu Cys Ile Met Gly  
20 25 30  
Asn Ser Gln Tyr Tyr Ile His Lys Ala Tyr His Arg Pro Lys His  
35 40 45  
Ile Gly His Asp Glu Trp Asp Val Ser Met Glu Arg Arg Asp Lys Lys

```

50 Val Val Glu Lys Ala Ala Ala Pro Ser Ser 345
65 (2) INFORMATION FOR SEQ ID NO:42:
70 (1) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 65 amino acids
  (B) TYPE: amino acid
  (C) STRANDEDNESS:
  (D) TOPOLOGY: linear
  (E) MOLECULE TYPE: peptide
  (F) FEATURE:
    (A) NAME/KEY: peptide
    (B) LOCATION: 1..65
  (D) OTHER INFORMATION: / Ceres Seq. ID 1011440
15 Met Ser Leu Val Trp Leu Glu Ala Met Leu Pro Leu Gly Ile Ile Gly
  5 10 15
  Gly Met Leu Ser Ile Met Gly Asn Ser Gln Tyr Tyr Ile His Lys Ala
  20 25 30
  Tyr His Gly Arg Pro Lys His Ile Gly His Asp Glu Trp Asp Val Ser
  35 40 45
  Met Glu Arg Arg Asp Lys Lys Val Val Glu Lys Ala Ala Pro Ser
  50 55 60
  Ser
  65 (2) INFORMATION FOR SEQ ID NO:43:
  (1) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 1011 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear
    (E) MOLECULE TYPE: DNA (genomic)
    (F) FEATURE:
      (A) NAME/KEY: -
      (B) LOCATION: 1..1011
      (D) OTHER INFORMATION: / Ceres Seq. ID 1011616
  (K) SEQUENCE DESCRIPTION: SEQ ID NO:43:
600 ctgttctt ttgtctctt agatctcc gtcgataac ccttcgcag gttcatlct
1200 ttctctga ttcttcaac aagtgctt taccagcat gcttcgcag cagatagct
1800 atccgcag ctgcagctg ttaagcgaa aattaaag gagggttaa gtttgaaa
2400 aactagtg ttgagctta aggaatac taagaact ttgcctgt accttggt-
3000 gggacaag gctcagacg gtccacat ttggagct ttccctct aaatcaaa
3600 tgaatttc cagctgcac ttctcaaa aaagttat ttgttggt tccccaca
4200 gctgattg agtcggga catcaatga gtagacat atgatatga agatttgt
4800 gtcgcctg caagcttc aactcgtt tcaatagc ctgttaac taattagt
5400 gttgagat ttgagtag ttggagaa tctctatc ctgaactg ttgtcttc
6000 gtttaact gaaatgca tcatcctc atcaattaa ggcataat aggttaagc
6600 ggtcagac ctgttgaa gatatgtt ggcatactg aagaagaaa gatcaagac
7200 cagatagc ctttcaaa ttgttcgt catgaaaa ctgatagt gttaccgtc
7800 aagcttgg aaaaatctt acagcttg cttaaatg gtttaaga gttgacatc
8400 aaacttga gtgacttg tccacaca atcccaag agttgaaga gttgagca
9000 tgaattcaat ccaagctag cccgaagat gataatcc latgaatg ctctcagt
9600 aaacacct aactcttag agtttagc tttaagtg ttcaagatt caaatggt
  (2) INFORMATION FOR SEQ ID NO:44:
  (1) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 252 amino acids
    (B) TYPE: amino acid
    (C) STRANDEDNESS:
    (D) TOPOLOGY: linear
    (E) MOLECULE TYPE: peptide
    (F) FEATURE:

```

	(A) NAME/KEY: peptide	346
	(B) LOCATION: 1..252	
	(D) OTHER INFORMATION: / Ceres Seq. ID I011617	
5	(X1) SEQUENCE DESCRIPTION: SEQ ID NO:44:	
1	Met Ser Ile Ser Gly Ala Ala Val Gly Ser Gly Arg Asn Leu Arg Arg	15
1	Ala Val Glu Phe Gly Lys Thr His Val Val Arg Pro Lys Gly Lys His	10
20	Gln Ala Thr Ile Val Trp Leu His Gly Leu Gly Asp Asn Gly Ser Ser	25
10	Trp Ser Gln Leu Leu Glu Thr Leu Pro Leu Pro Asn Ile Lys Trp Ile	40
35	Cys Pro Thr Ala Pro Ser Gln Pro Ile Ser Leu Phe Gly Gly Phe Pro	55
15	Ser Thr Ala Trp Phe Asp Val Val Asp Ile Asn Glu Asp Gly Pro Asp	70
65	Asp Met Glu Gly Leu Asp Val Ala Ala His Val Ala Asn Leu Leu	85
20	Ser Asn Glu Pro Ala Asp Ile Lys Leu Gly Val Gly Gly Phe Ser Met	100
115	Gly Ala Ala Thr Ser Leu Tyr Ser Ala Thr Cys Phe Ala Leu Gly Lys	115
25	Leu Ser Gly Trp Leu Pro Cys Ala Lys Thr Leu Ala Gly Lys Leu Glu	130
145	Glu Glu Gln Ile Lys Asn Arg Ala Ala Ser Leu Pro Ile Val Val Cys	145
30	His Gly Lys Ala Asp Asp Val Val Pro Phe Lys Phe Gly Glu Lys Ser	160
180	Ser Gln Ala Leu Leu Ser Asn Gly Phe Lys Lys Val Thr Phe Lys Pro	175
35	Tyr Ser Ala Leu Gly His His Thr Ile Pro Gln Glu Leu Asp Glu Leu	190
225	Cys Ala Trp Leu Thr Ser Thr Leu Ser Leu Glu Gly	205
245		220
	(2) INFORMATION FOR SEQ ID NO:45:	
	(1) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 155 amino acids	
	(B) TYPE: amino acid	
	(C) STRANDNESS:	
	(D) TOPOLOGY: linear	
	(11) MOLECULE TYPE: peptide	
	(1X) FEATURE:	
	(A) NAME/KEY: peptide	
	(B) LOCATION: 1..155	
	(D) OTHER INFORMATION: / Ceres Seq. ID I011618	
50	(X1) SEQUENCE DESCRIPTION: SEQ ID NO:45:	
1	Met Gln Gly Leu Asp Val Ala Ala His Val Ala Asn Leu Leu Ser	15
5	Asn Glu Pro Ala Asp Ile Lys Leu Gly Val Gly Gly Phe Ser Met Gly	10
20	Ala Ala Thr Ser Leu Tyr Ser Ala Thr Cys Phe Ala Leu Gly Lys Tyr	25
35	Gly Asn Gly Asn Pro Tyr Pro Ile Asn Leu Ser Ala Ile Ile Gly Leu	40
55	Ser Gly Trp Leu Pro Cys Ala Lys Thr Leu Ala Gly Lys Leu Glu Gln	55
65	Glu Gln Ile Lys Asn Arg Ala Ala Ser Leu Pro Ile Val Val Cys His	70
85	Gly Lys Ala Asp Asp Val Val Pro Phe Lys Phe Gly Glu Lys Ser Ser	90
100		95
		110

347

Gln Ala Leu Leu Ser Asn Gly Phe Lys Lys Val Thr Phe Lys Pro Tyr  
115 120 125  
Ser Ala Leu Gly His Thr Ile Pro Gln Glu Leu Asp Glu Leu Cys  
130 135 140 145 150 155

5 Ala Trp Leu Thr Ser Thr Leu Ser Leu Glu Gly  
145 150 155

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 543 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..643

(D) OTHER INFORMATION: / Ceres Seq. ID 1011631

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

agaatgtgat ggaatggga gtaataaga tgaatgtgg gattaagaa gagaagaag  
120  
agaaggtctt tatgttaga ttgacatgt aat-cagtcat ttcacattgg ggaagccaa  
180  
ggaaccatg gacggcggg gggccatcg aatagacct cgaatggtt tgttcccaa  
240  
ccatacat ggaatgaagt ggaagagag ccatcatca caaccattc cgcggctag  
300  
ggtacacct aggaatgctt gggatggg gaagagagc taggtgttca agatccgag  
360  
agaaagay gacaaggttg ttctccaag agataagta cgaatagct aatataag  
420  
cagaagaa gctccgatg aaagaaggt tgcacaag atctcaatt ggtgtgctc  
480  
actaagaac taatcaatt aggaataa aattacttg cctcattt ctttttctt  
540  
tgctaatg ttgtgtatg ttgattct tttctgat tcatagaga ttgtgacgt  
600  
ctttg-gagc tagctatga cataaatca tcaccaaaa aa-gtgacta tcttgaaac  
actaatat atagtcata taagtgaat t-tgattgt ggc

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 140 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptido

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..140

(D) OTHER INFORMATION: / Ceres Seq. ID 1011632

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Asn Val Met Glu Met Gly Val Asn Glu Met Ser Gly Gly Ile Lys Glu  
1 5 10 15  
Glu Lys Lys Glu Lys Ala Leu Met Leu Arg Leu Asp Tyr Glu Ser Val  
20 25 30 35  
Ile Ser Thr Trp Gly Gly Gln Gly Ile Pro Trp Thr Ala Arg Val Pro  
40 45  
Ser Glu Ile Asp Leu Asp Met Val Cys Phe Pro Thr His Thr Met Gly  
50 55 60  
Glu Ser Gly Ala Glu Ala His His Asn His Phe Arg Gly Leu Gly  
65 70 75 80  
Leu His Leu Gly Asp Ala Gly Asp Gly Gly Arg Glu Ala Arg Val Ser  
85 90 95  
Arg Tyr Arg Glu Lys Arg Arg Thr Arg Leu Phe Ser Lys Lys Ile Arg  
100 105 110 115  
Tyr Glu Val Arg Lys Leu Asn Ala Glu Lys Arg Pro Arg Met Lys Gly  
120 125  
Arg Phe Val Lys Arg Ser Ser Ile Gly Val Ala His  
130 135 140

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 138 amino acids

348

(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..138

(D) OTHER INFORMATION: / Ceres Seq. ID 1011633

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Met Glu Met Gly Val Asn Glu Met Ser Gly Gly Ile Lys Glu Glu Lys  
1 5 10 15  
Lys Glu Lys Ala Leu Met Leu Arg Leu Asp Tyr Glu Ser Val Ile Ser  
20 25 30  
Thr Trp Gly Gly Gln Gly Ile Pro Trp Thr Ala Arg Val Pro Ser Glu  
35 40 45  
Ile Asp Leu Asp Met Val Cys Pro Thr His Thr Met Gly Glu Ser  
50 55 60  
Gly Ala Glu Ala His His Asn His Phe Arg Gly Leu Gly Leu His  
65 70 75 80  
Leu Gly Asp Ala Gly Asp Gly Arg Glu Ala Arg Val Ser Arg Tyr  
85 90 95  
Arg Glu Lys Arg Arg Thr Arg Leu Phe Ser Lys Ile Arg Tyr Glu  
100 105 110  
Val Arg Lys Leu Asn Ala Glu Lys Arg Pro Arg Met Lys Gly Arg Phe  
115 120 125  
Val Lys Arg Ser Ser Ile Gly Val Ala His  
130 135

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 136 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..136

(D) OTHER INFORMATION: / Ceres Seq. ID 1011634

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

Met Gly Val Asn Glu Met Ser Gly Gly Ile Lys Glu Glu Lys Lys Glu  
1 5 10 15  
Lys Ala Leu Met Leu Arg Leu Asp Tyr Glu Ser Val Ile Ser Thr Trp  
20 25 30  
Gly Gly Gln Gly Ile Pro Trp Thr Ala Arg Val Pro Ser Glu Ile Asp  
35 40 45  
Leu Asp Met Val Cys Phe Pro Thr His Thr Met Gly Glu Ser Gly Ala  
50 55 60  
Glu Ala His His Asn His Phe Arg Gly Leu Gly Leu His Leu Gly  
65 70 75 80  
Asp Ala Gly Asp Gly Arg Glu Ala Arg Val Ser Arg Tyr Arg Glu  
85 90 95  
Lys Arg Arg Thr Arg Leu Phe Ser Lys Lys Ile Arg Tyr Glu Val Arg  
100 105 110  
Lys Leu Asn Ala Glu Lys Arg Pro Arg Met Lys Gly Arg Phe Val Lys  
115 120 125  
Arg Ser Ser Ile Gly Val Ala His  
130 135

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 663 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

349

(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

(A) NAME/KEY: -  
(B) LOCATION: 1..663

(D) OTHER INFORMATION: / Ceres Seq. ID 1011714

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

gtgttgata ataatgaga cagagaagc gatacaaga gtcgcgtcg tcgtttcat 60  
tttcaacgga acctcaatt tatagtagc ccccggtcc tcaacgga acctcaatt 120  
cgctctccc gggccacc tcgattcgg agagagctt gaagatggg cagcagaga 180  
agtatggg gaacagcgc taagattga aaagatgaag cttttgactg ttcaaacaa 240  
tgtttcaa gaagacaca cgcacaca ctactctct gtttcgatac gtcgaggtt 300  
ggtgatac agtcaagac cgaagatc ggaacagag aagtggaag gtcgaggtt 360  
gtatgattg gagaatrac caaaccttt gtttggcca ctgagaaat ttttggagg 420  
tggttcaat ccttcaact atgtgtgtg agactaatg atgaagagt taatgattga 480  
tttgattg aatgttcaa atgttgcat tggtagtag gtaggattt cgttaggg- 540  
gcgagagtc cagattcaa ttctcgaat gccctctct ttactcttt ttattgtat- 600  
ttttgttcc tttttctta attagtaata aacttatac gttgcctatc gattttatt- 660  
ctc

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 151 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..151

(D) OTHER INFORMATION: / Ceres Seq. ID 1011715

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

Cys Ser Ile Ile Met Ser Thr Gly Glu Ala Ile Pro Arg Val Ala Val 15  
1 Val Val Phe Ile Leu Asn Gly Asn Ser Ile Leu Leu Gly Arg Arg Arg 20  
Ser Ser Ile Gly Asn Ser Thr Phe Ala Leu Pro Gly Gly His Leu Glu 30  
35 Phe Gly Glu Ser Phe Glu Glu Cys Ala Ala Arg Glu Val Met Glu Glu 40  
Thr Gly Leu Lys Ile Glu Lys Met Lys Leu Leu Thr Val Thr Asn Asn 50  
63 Val Phe Lys Glu Ala Pro Thr Pro Ser His Tyr Val Ser Val Ile 60  
Arg Ala Val Leu Val Asp Pro Ser Glu Glu Pro Lys Asn Met Glu Pro 70  
85 Glu Lys Cys Glu Gly Trp Asp Trp Tyr Asp Trp Glu Asn Leu Pro Lys 90  
100 Pro Leu Phe Trp Pro Leu Glu Lys Leu Phe Gly Ser Gly Phe Asn Pro 110  
120 Phe Thr His Gly Gly Gly Asp 130  
145

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 147 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..147

(D) OTHER INFORMATION: / Ceres Seq. ID 1011716

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

Met Ser Thr Gly Glu Ala Ile Pro Arg Val 10  
1 Leu Asn Gly Asn Ser Ile Leu Leu Gly Arg Arg 20  
25 Asn Ser Thr Phe Ala Leu Pro Gly Gly His Leu G. 30

35 Phe Glu Glu Cys Ala Ala Arg Glu Val Met Glu Glu Thr. 40  
50 Ile Glu Lys Met Lys Leu Leu Thr Val Thr Asn Asn Val Phe L, 55  
65 A.A Pro Thr Pro Ser His Tyr Val Ser Val Ser Ile Arg Ala Val Leu 60  
70 Val Asp Pro Ser Glu Glu Pro Lys Asn Met Glu Pro Glu Lys Cys Glu 75  
85 Gly Trp Asp Trp Tyr Asp Trp Glu Asn Leu Pro Lys Pro Leu Phe Trp 80  
90 Pro Leu Glu Lys Leu Phe Gly Ser Gly Phe Asn Pro Phe Thr His Gly 95  
100 Gly Gly Asp 105  
110

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

35

351  
 acagcttgc agattcttc cccagctgt taagaaatg ttaactaga ttaattagtg 360  
 ccagcttgc cagattatg tggtaacg tggcttatc ttgctcagta ctacttaaa 420  
 taataatga gtagatga ttgatgata tcaatcat gttgaatg: tatgttttc 480  
 acctataaa acatgaara tactttaga atagcttgc tegt

5 (2) INFORMATION FOR SEQ ID NO:55:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 110 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

(A) NAME/KEY: peptide  
 (B) LOCATION: 1..110

(D) OTHER INFORMATION: / Ceres Seq. ID 1011785  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:  
 Lys Thr Glu Gln Thr Gln Thr Gln Thr Glu Asn Gln Arg Arg Arg Lys 1  
 Met Val Lys Val Met Trp Val Ser Val Leu Ala Leu Ala Ala Ile 15  
 Leu Leu Leu Thr Val Pro Val Ala Glu Gly Val Thr Cys Ser Pro Met: 20  
 Gln Leu Ala Ser Cys Ala Ala Ala Met Thr Ser Ser Ser Pro Ser 25  
 Glu Ala Cys Cys Thr Lys Leu Arg Glu Gln Gln Pro Cys Leu Cys Gly 30  
 Tyr Met Arg Asn Pro Thr Leu Arg Gln Tyr Val Ser Ser Pro Asn Ala 35  
 Arg Lys Val Ser Asn Ser Cys Lys Ile Pro Ser Pro Ser Cys 40

10 (2) INFORMATION FOR SEQ ID NO:56:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 94 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

15 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..94  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011786  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:  
 Met Val Lys Val Met Trp Val Ser Val Leu Ala Leu Ala Ala Ile 1  
 Leu Leu Leu Thr Val Pro Val Ala Glu Gly Val Thr Cys Ser Pro Met: 15  
 Gln Leu Ala Ser Cys Ala Ala Ala Met Thr Ser Ser Ser Pro Ser 20  
 Glu Ala Cys Cys Thr Lys Leu Arg Glu Gln Gln Pro Cys Leu Cys Gly 25  
 Tyr Met Arg Asn Pro Thr Leu Arg Gln Tyr Val Ser Ser Pro Asn Ala 30  
 Arg Lys Val Ser Asn Ser Cys Lys Ile Pro Ser Pro Ser Cys 35

20 (2) INFORMATION FOR SEQ ID NO:57:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 110 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

25 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..110  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011787  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:  
 Lys Thr Glu Gln Thr Gln Thr Gln Thr Glu Asn Gln Arg Arg Arg Lys 1  
 Met Val Lys Val Met Trp Val Ser Val Leu Ala Leu Ala Ala Ile 15  
 Leu Leu Leu Thr Val Pro Val Ala Glu Gly Val Thr Cys Ser Pro Met: 20  
 Gln Leu Ala Ser Cys Ala Ala Ala Met Thr Ser Ser Ser Pro Ser 25  
 Glu Ala Cys Cys Thr Lys Leu Arg Glu Gln Gln Pro Cys Leu Cys Gly 30  
 Tyr Met Arg Asn Pro Thr Leu Arg Gln Tyr Val Ser Ser Pro Asn Ala 35  
 Arg Lys Val Ser Asn Ser Cys Lys Ile Pro Ser Pro Ser Cys 40

30 (2) INFORMATION FOR SEQ ID NO:58:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 107 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:

35 (A) NAME/KEY: -  
 (B) LOCATION: 1..1007  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011820  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:  
 ccaaatctct ccagtagcg tctttgaat ctctcgatt tagcttagg gttcgtagag 60  
 cagctagcta ttcacagcc actaatgga aaagaaatc taatagaaga ctgttagag 120  
 acagaggtt cctttgact tcacacagta atgaacgtg tctctctc taactctcc 180  
 gacgactta tctttctgc ttttgcgc cctttctc attatgacg aacttgcgt 240  
 aaactggat gtaatttag taccataag tgaataac gatgggtgt cttagcatg 300  
 cctatggtta agtttcaga tgggttcaa atggaagaa atggagtc agttgcgaa 360  
 ctattcggg agatacaga aaccttga agctacaga gtgaacag tgttaacata 420  
 actgtatgg tgaatgacat gctctgctg gaaatgcta ctacggcg caactcagt 480  
 cagctatgg acttttgc ttttgcac acataagtt ctgaagcaa ttgtcattg 540  
 gtcctctca atcagaaga tatacggc agcatggga gactgcatt ttgtctacg 600  
 atggtatgcc ttgcagatgt ttgtatag gcagagcct tagctctgg tttagcaat 660  
 gatgacatg gccaatgac tttctgaac aaaggataa gcaactcag tagagaagc 720  
 tctgagaaca agttcagaa ttctcaatt agatcaagg aaatggtat cgactattc 780  
 tatcttggt gcagaagctg agcatagac accgtttga gcatagcat ggtatgaga 840  
 tgcattaga cagttttgt attgcgctt aggttaaca gagacactt gttaacatt 900  
 ttctcttag tctctctct ggttgaact tgaactta gctaaagtt aatgcttag 960  
 acacagaac tttagacat ggaatagg taacttcta taacttcta tatcttt

40 (2) INFORMATION FOR SEQ ID NO:59:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 266 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

45 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..266  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011821  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:  
 Pro Asn Phe Ser Met Asp Arg Ser Leu Asn Leu Leu Asp Leu Ala Leu 1  
 Gly Phe Asp Glu Gln Leu Ala Ile Pro Ser Pro Leu Asn Gly Lys Val: 15  
 Ile Leu Ile Glu Asp Cys Val Glu Thr Ser Gly Ser Phe Val Leu His 30

50 (2) INFORMATION FOR SEQ ID NO:60:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 90 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

55 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..266  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011822  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:  
 Pro Asn Phe Ser Met Asp Arg Ser Leu Asn Leu Leu Asp Leu Ala Leu 1  
 Gly Phe Asp Glu Gln Leu Ala Ile Pro Ser Pro Leu Asn Gly Lys Val: 15  
 Ile Leu Ile Glu Asp Cys Val Glu Thr Ser Gly Ser Phe Val Leu His 30

60 (2) INFORMATION FOR SEQ ID NO:61:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 107 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:

65 (A) NAME/KEY: -  
 (B) LOCATION: 1..1007  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011823  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:  
 ccaaatctct ccagtagcg tctttgaat ctctcgatt tagcttagg gttcgtagag 60  
 cagctagcta ttcacagcc actaatgga aaagaaatc taatagaaga ctgttagag 120  
 acagaggtt cctttgact tcacacagta atgaacgtg tctctctc taactctcc 180  
 gacgactta tctttctgc ttttgcgc cctttctc attatgacg aacttgcgt 240  
 aaactggat gtaatttag taccataag tgaataac gatgggtgt cttagcatg 300  
 cctatggtta agtttcaga tgggttcaa atggaagaa atggagtc agttgcgaa 360  
 ctattcggg agatacaga aaccttga agctacaga gtgaacag tgttaacata 420  
 actgtatgg tgaatgacat gctctgctg gaaatgcta ctacggcg caactcagt 480  
 cagctatgg acttttgc ttttgcac acataagtt ctgaagcaa ttgtcattg 540  
 gtcctctca atcagaaga tatacggc agcatggga gactgcatt ttgtctacg 600  
 atggtatgcc ttgcagatgt ttgtatag gcagagcct tagctctgg tttagcaat 660  
 gatgacatg gccaatgac tttctgaac aaaggataa gcaactcag tagagaagc 720  
 tctgagaaca agttcagaa ttctcaatt agatcaagg aaatggtat cgactattc 780  
 tatcttggt gcagaagctg agcatagac accgtttga gcatagcat ggtatgaga 840  
 tgcattaga cagttttgt attgcgctt aggttaaca gagacactt gttaacatt 900  
 ttctcttag tctctctct ggttgaact tgaactta gctaaagtt aatgcttag 960  
 acacagaac tttagacat ggaatagg taacttcta taacttcta tatcttt

70 (2) INFORMATION FOR SEQ ID NO:62:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 266 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

75 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..266  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011824  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:  
 Pro Asn Phe Ser Met Asp Arg Ser Leu Asn Leu Leu Asp Leu Ala Leu 1  
 Gly Phe Asp Glu Gln Leu Ala Ile Pro Ser Pro Leu Asn Gly Lys Val: 15  
 Ile Leu Ile Glu Asp Cys Val Glu Thr Ser Gly Ser Phe Val Leu His 30

80 (2) INFORMATION FOR SEQ ID NO:63:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 107 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:

85 (A) NAME/KEY: -  
 (B) LOCATION: 1..1007  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011825  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:  
 ccaaatctct ccagtagcg tctttgaat ctctcgatt tagcttagg gttcgtagag 60  
 cagctagcta ttcacagcc actaatgga aaagaaatc taatagaaga ctgttagag 120  
 acagaggtt cctttgact tcacacagta atgaacgtg tctctctc taactctcc 180  
 gacgactta tctttctgc ttttgcgc cctttctc attatgacg aacttgcgt 240  
 aaactggat gtaatttag taccataag tgaataac gatgggtgt cttagcatg 300  
 cctatggtta agtttcaga tgggttcaa atggaagaa atggagtc agttgcgaa 360  
 ctattcggg agatacaga aaccttga agctacaga gtgaacag tgttaacata 420  
 actgtatgg tgaatgacat gctctgctg gaaatgcta ctacggcg caactcagt 480  
 cagctatgg acttttgc ttttgcac acataagtt ctgaagcaa ttgtcattg 540  
 gtcctctca atcagaaga tatacggc agcatggga gactgcatt ttgtctacg 600  
 atggtatgcc ttgcagatgt ttgtatag gcagagcct tagctctgg tttagcaat 660  
 gatgacatg gccaatgac tttctgaac aaaggataa gcaactcag tagagaagc 720  
 tctgagaaca agttcagaa ttctcaatt agatcaagg aaatggtat cgactattc 780  
 tatcttggt gcagaagctg agcatagac accgtttga gcatagcat ggtatgaga 840  
 tgcattaga cagttttgt attgcgctt aggttaaca gagacactt gttaacatt 900  
 ttctcttag tctctctct ggttgaact tgaactta gctaaagtt aatgcttag 960  
 acacagaac tttagacat ggaatagg taacttcta taacttcta tatcttt

90 (2) INFORMATION FOR SEQ ID NO:64:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 266 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

95 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..266  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011826  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:  
 Pro Asn Phe Ser Met Asp Arg Ser Leu Asn Leu Leu Asp Leu Ala Leu 1  
 Gly Phe Asp Glu Gln Leu Ala Ile Pro Ser Pro Leu Asn Gly Lys Val: 15  
 Ile Leu Ile Glu Asp Cys Val Glu Thr Ser Gly Ser Phe Val Leu His 30

100 (2) INFORMATION FOR SEQ ID NO:65:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 107 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:

105 (A) NAME/KEY: -  
 (B) LOCATION: 1..1007  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011827  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:  
 ccaaatctct ccagtagcg tctttgaat ctctcgatt tagcttagg gttcgtagag 60  
 cagctagcta ttcacagcc actaatgga aaagaaatc taatagaaga ctgttagag 120  
 acagaggtt cctttgact tcacacagta atgaacgtg tctctctc taactctcc 180  
 gacgactta tctttctgc ttttgcgc cctttctc attatgacg aacttgcgt 240  
 aaactggat gtaatttag taccataag tgaataac gatgggtgt cttagcatg 300  
 cctatggtta agtttcaga tgggttcaa atggaagaa atggagtc agttgcgaa 360  
 ctattcggg agatacaga aaccttga agctacaga gtgaacag tgttaacata 420  
 actgtatgg tgaatgacat gctctgctg gaaatgcta ctacggcg caactcagt 480  
 cagctatgg acttttgc ttttgcac acataagtt ctgaagcaa ttgtcattg 540  
 gtcctctca atcagaaga tatacggc agcatggga gactgcatt ttgtctacg 600  
 atggtatgcc ttgcagatgt ttgtatag gcagagcct tagctctgg tttagcaat 660  
 gatgacatg gccaatgac tttctgaac aaaggataa gcaactcag tagagaagc 720  
 tctgagaaca agttcagaa ttctcaatt agatcaagg aaatggtat cgactattc 780  
 tatcttggt gcagaagctg agcatagac accgtttga gcatagcat ggtatgaga 840  
 tgcattaga cagttttgt attgcgctt aggttaaca gagacactt gttaacatt 900  
 ttctcttag tctctctct ggttgaact tgaactta gctaaagtt aatgcttag 960  
 acacagaac tttagacat ggaatagg taacttcta taacttcta tatcttt

110 (2) INFORMATION FOR SEQ ID NO:66:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 266 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

115 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..266  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011828  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:  
 Pro Asn Phe Ser Met Asp Arg Ser Leu Asn Leu Leu Asp Leu Ala Leu 1  
 Gly Phe Asp Glu Gln Leu Ala Ile Pro Ser Pro Leu Asn Gly Lys Val: 15  
 Ile Leu Ile Glu Asp Cys Val Glu Thr Ser Gly Ser Phe Val Leu His 30

120 (2) INFORMATION FOR SEQ ID NO:67:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 107 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:

125 (A) NAME/KEY: -  
 (B) LOCATION: 1..1007  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011829  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:  
 ccaaatctct ccagtagcg tctttgaat ctctcgatt tagcttagg gttcgtagag 60  
 cagctagcta ttcacagcc actaatgga aaagaaatc taatagaaga ctgttagag 120  
 acagaggtt cctttgact tcacacagta atgaacgtg tctctctc taactctcc 180  
 gacgactta tctttctgc ttttgcgc cctttctc attatgacg aacttgcgt 240  
 aaactggat gtaatttag taccataag tgaataac gatgggtgt cttagcatg 300  
 cctatggtta agtttcaga tgggttcaa atggaagaa atggagtc agttgcgaa 360  
 ctattcggg agatacaga aaccttga agctacaga gtgaacag tgttaacata 420  
 actgtatgg tgaatgacat gctctgctg gaaatgcta ctacggcg caactcagt 480  
 cagctatgg acttttgc ttttgcac acataagtt ctgaagcaa ttgtcattg 540  
 gtcctctca atcagaaga tatacggc agcatggga gactgcatt ttgtctacg 600  
 atggtatgcc ttgcagatgt ttgtatag gcagagcct tagctctgg tttagcaat 660  
 gatgacatg gccaatgac tttctgaac aaaggataa gcaactcag tagagaagc 720  
 tctgagaaca agttcagaa ttctcaatt agatcaagg aaatggtat cgactattc 780  
 tatcttggt gcagaagctg agcatagac accgtttga gcatagcat ggtatgaga 840  
 tgcattaga cagttttgt attgcgctt aggttaaca gagacactt gttaacatt 900  
 ttctcttag tctctctct ggttgaact tgaactta gctaaagtt aatgcttag 960  
 acacagaac tttagacat ggaatagg taacttcta taacttcta tatcttt

130 (2) INFORMATION FOR SEQ ID NO:68:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 266 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

135 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..266  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011830  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:  
 Pro Asn Phe Ser Met Asp Arg Ser Leu Asn Leu Leu Asp Leu Ala Leu 1  
 Gly Phe Asp Glu Gln Leu Ala Ile Pro Ser Pro Leu Asn Gly Lys Val: 15  
 Ile Leu Ile Glu Asp Cys Val Glu Thr Ser Gly Ser Phe Val Leu His 30

353  
 35 Gln Leu Met Lys Arg Val Leu Ser Ser Asn Ser Ser Asp Ala Leu Ile 45  
 50 Phe Leu Ala Phe Ala Arg Pro Phe Ser His Tyr Asp Arg Ile Leu Arg 60  
 65 Lys Leu Gly Cys Asn Leu Ala Thr His Lys Ser Asn Asn Arg Leu Val 75  
 85 Phe Asp Met Leu Met Val Lys Cys Ser Asp Gly Asp Gln Met Glu 95  
 100 Asp Asn Val Ser Ala Val Ala Lys Leu Phe Arg Gln Ile Gln Glu Thr 110  
 115 Val Arg Lys Leu Gln Ser Val Thr Ser Gly Asn Ile Thr Val Met Val 125  
 130 Asp Asp Met Ser Leu Leu Glu Ile Ala Thr Thr Gly Ser Asn Ser Asp 140  
 145 His Val Leu Asp Phe Leu His Tyr Cys His Thr Leu Ser Ser Glu Ser 155  
 165 Asn Cys Ser Leu Val Ile Leu Asn His Glu Asp Ile Tyr Ala Ser Met 175  
 185 Glu Arg Pro Ala Phe Leu Leu Gln Met Val Cys Leu Ala Asp Val Val 190  
 195 Ile Lys Ala Glu Pro Leu Ala Ser Gly Leu Ala Asn Asp Val Val 205  
 210 Gln Leu Thr Val Leu Asn Lys Gly Ile Ser Asn Ser Gly Arg Gly Ser 220  
 225 Ser Arg Asn Lys Leu Gln Asn Phe Gln Phe Arg Ile Lys Glu Asn Gly 235  
 245 Ile Asp Tyr Phe Tyr Pro Gly Cys Arg Ser 250  
 260  
 265  
 (2) INFORMATION FOR SEQ ID NO:60:  
 (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 262 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..262  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011822  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:  
 Met Asp Arg Ser Leu Asn Leu Leu Asp Leu Leu Ala Leu Gly Phe Asp Glu 1  
 5 10 15  
 Cln Leu Ala Ile Pro Ser Pro Leu Asn Gly Lys Val Ile Leu Ile Glu 20  
 25 30  
 Asp Cys Val Glu Thr Ser Gly Ser Phe Val Leu His Gln Leu Met Lys 35  
 40 45  
 Arg Val Leu Ser Ser Asn Ser Asp Ala Leu Ile Phe Leu Ala Phe 50  
 55 60  
 Ala Arg Pro Phe Ser His Tyr Asp Arg Ile Leu Arg Lys Leu Gly Cys 65  
 70 75  
 Asn Leu Ala Thr His Lys Ser Asn Asn Arg Leu Val Phe Asp Met 80  
 85 90  
 Leu Met Val Lys Cys Ser Asp Gly Asp Gln Met Glu Asp Asn Val Ser 100  
 105 110  
 Ala Val Ala Lys Leu Phe Arg Glu Ile Gln Glu Thr Val Arg Lys Leu 115  
 120 125  
 Gln Ser Val Thr Ser Gly Asn Ile Thr Val Met Val Asp Asp Met Ser 130  
 135 140  
 Leu Leu Glu Ile Ala Thr Thr Gly Ser Asn Ser Asp His Val Leu Asp 145  
 150 155  
 Phe Leu His Tyr Cys His Thr Leu Ser Ser Glu Ser Asn Cys Ser Leu 160  
 165 170

354  
 170  
 175  
 180  
 185  
 190  
 195  
 200  
 205  
 210  
 215  
 220  
 225  
 230  
 235  
 240  
 245  
 250  
 255  
 260  
 265  
 270  
 275  
 280  
 285  
 290  
 295  
 300  
 305  
 310  
 315  
 320  
 325  
 330  
 335  
 340  
 345  
 350  
 355  
 360  
 365  
 370  
 375  
 380  
 385  
 390  
 395  
 400  
 405  
 410  
 415  
 420  
 425  
 430  
 435  
 440  
 445  
 450  
 455  
 460  
 465  
 470  
 475  
 480  
 485  
 490  
 495  
 500  
 505  
 510  
 515  
 520  
 525  
 530  
 535  
 540  
 545  
 550  
 555  
 560  
 565  
 570  
 575  
 580  
 585  
 590  
 595  
 600  
 605  
 610  
 615  
 620  
 625  
 630  
 635  
 640  
 645  
 650  
 655  
 660  
 665  
 670  
 675  
 680  
 685  
 690  
 695  
 700  
 705  
 710  
 715  
 720  
 725  
 730  
 735  
 740  
 745  
 750  
 755  
 760  
 765  
 770  
 775  
 780  
 785  
 790  
 795  
 800  
 805  
 810  
 815  
 820  
 825  
 830  
 835  
 840  
 845  
 850  
 855  
 860  
 865  
 870  
 875  
 880  
 885  
 890  
 895  
 900  
 905  
 910  
 915  
 920  
 925  
 930  
 935  
 940  
 945  
 950  
 955  
 960  
 965  
 970  
 975  
 980  
 985  
 990  
 995  
 1000  
 1005  
 1010  
 1015  
 1020  
 1025  
 1030  
 1035  
 1040  
 1045  
 1050  
 1055  
 1060  
 1065  
 1070  
 1075  
 1080  
 1085  
 1090  
 1095  
 1100  
 1105  
 1110  
 1115  
 1120  
 1125  
 1130  
 1135  
 1140  
 1145  
 1150  
 1155  
 1160  
 1165  
 1170  
 1175  
 1180  
 1185  
 1190  
 1195  
 1200  
 1205  
 1210  
 1215  
 1220  
 1225  
 1230  
 1235  
 1240  
 1245  
 1250  
 1255  
 1260  
 1265  
 1270  
 1275  
 1280  
 1285  
 1290  
 1295  
 1300  
 1305  
 1310  
 1315  
 1320  
 1325  
 1330  
 1335  
 1340  
 1345  
 1350  
 1355  
 1360  
 1365  
 1370  
 1375  
 1380  
 1385  
 1390  
 1395  
 1400  
 1405  
 1410  
 1415  
 1420  
 1425  
 1430  
 1435  
 1440  
 1445  
 1450  
 1455  
 1460  
 1465  
 1470  
 1475  
 1480  
 1485  
 1490  
 1495  
 1500  
 1505  
 1510  
 1515  
 1520  
 1525  
 1530  
 1535  
 1540  
 1545  
 1550  
 1555  
 1560  
 1565  
 1570  
 1575  
 1580  
 1585  
 1590  
 1595  
 1600  
 1605  
 1610  
 1615  
 1620  
 1625  
 1630  
 1635  
 1640  
 1645  
 1650  
 1655  
 1660  
 1665  
 1670  
 1675  
 1680  
 1685  
 1690  
 1695  
 1700  
 1705  
 1710  
 1715  
 1720  
 1725  
 1730  
 1735  
 1740  
 1745  
 1750  
 1755  
 1760  
 1765  
 1770  
 1775  
 1780  
 1785  
 1790  
 1795  
 1800  
 1805  
 1810  
 1815  
 1820  
 1825  
 1830  
 1835  
 1840  
 1845  
 1850  
 1855  
 1860  
 1865  
 1870  
 1875  
 1880  
 1885  
 1890  
 1895  
 1900  
 1905  
 1910  
 1915  
 1920  
 1925  
 1930  
 1935  
 1940  
 1945  
 1950  
 1955  
 1960  
 1965  
 1970  
 1975  
 1980  
 1985  
 1990  
 1995  
 2000  
 2005  
 2010  
 2015  
 2020  
 2025  
 2030  
 2035  
 2040  
 2045  
 2050  
 2055  
 2060  
 2065  
 2070  
 2075  
 2080  
 2085  
 2090  
 2095  
 2100  
 2105  
 2110  
 2115  
 2120  
 2125  
 2130  
 2135  
 2140  
 2145  
 2150  
 2155  
 2160  
 2165  
 2170  
 2175  
 2180  
 2185  
 2190  
 2195  
 2200  
 2205  
 2210  
 2215  
 2220  
 2225  
 2230  
 2235  
 2240  
 2245  
 2250  
 2255  
 2260  
 2265  
 2270  
 2275  
 2280  
 2285  
 2290  
 2295  
 2300  
 2305  
 2310  
 2315  
 2320  
 2325  
 2330  
 2335  
 2340  
 2345  
 2350  
 2355  
 2360  
 2365  
 2370  
 2375  
 2380  
 2385  
 2390  
 2395  
 2400  
 2405  
 2410  
 2415  
 2420  
 2425  
 2430  
 2435  
 2440  
 2445  
 2450  
 2455  
 2460  
 2465  
 2470  
 2475  
 2480  
 2485  
 2490  
 2495  
 2500  
 2505  
 2510  
 2515  
 2520  
 2525  
 2530  
 2535  
 2540  
 2545  
 2550  
 2555  
 2560  
 2565  
 2570  
 2575  
 2580  
 2585  
 2590  
 2595  
 2600  
 2605  
 2610  
 2615  
 2620  
 2625  
 2630  
 2635  
 2640  
 2645  
 2650  
 2655  
 2660  
 2665  
 2670  
 2675  
 2680  
 2685  
 2690  
 2695  
 2700  
 2705  
 2710  
 2715  
 2720  
 2725  
 2730  
 2735  
 2740  
 2745  
 2750  
 2755  
 2760  
 2765  
 2770  
 2775  
 2780  
 2785  
 2790  
 2795  
 2800  
 2805  
 2810  
 2815  
 2820  
 2825  
 2830  
 2835  
 2840  
 2845  
 2850  
 2855  
 2860  
 2865  
 2870  
 2875  
 2880  
 2885  
 2890  
 2895  
 2900  
 2905  
 2910  
 2915  
 2920  
 2925  
 2930  
 2935  
 2940  
 2945  
 2950  
 2955  
 2960  
 2965  
 2970  
 2975  
 2980  
 2985  
 2990  
 2995  
 3000  
 3005  
 3010  
 3015  
 3020  
 3025  
 3030  
 3035  
 3040  
 3045  
 3050  
 3055  
 3060  
 3065  
 3070  
 3075  
 3080  
 3085  
 3090  
 3095  
 3100  
 3105  
 3110  
 3115  
 3120  
 3125  
 3130  
 3135  
 3140  
 3145  
 3150  
 3155  
 3160  
 3165  
 3170  
 3175  
 3180  
 3185  
 3190  
 3195  
 3200  
 3205  
 3210  
 3215  
 3220  
 3225  
 3230  
 3235  
 3240  
 3245  
 3250  
 3255  
 3260  
 3265  
 3270  
 3275  
 3280  
 3285  
 3290  
 3295  
 3300  
 3305  
 3310  
 3315  
 3320  
 3325  
 3330  
 3335  
 3340  
 3345  
 3350  
 3355  
 3360  
 3365  
 3370  
 3375  
 3380  
 3385  
 3390  
 3395  
 3400  
 3405  
 3410  
 3415  
 3420  
 3425  
 3430  
 3435  
 3440  
 3445  
 3450  
 3455  
 3460  
 3465  
 3470  
 3475  
 3480  
 3485  
 3490  
 3495  
 3500  
 3505  
 3510  
 3515  
 3520  
 3525  
 3530  
 3535  
 3540  
 3545  
 3550  
 3555  
 3560  
 3565  
 3570  
 3575  
 3580  
 3585  
 3590  
 3595  
 3600  
 3605  
 3610  
 3615  
 3620  
 3625  
 3630  
 3635  
 3640  
 3645  
 3650  
 3655  
 3660  
 3665  
 3670  
 3675  
 3680  
 3685  
 3690  
 3695  
 3700  
 3705  
 3710  
 3715  
 3720  
 3725  
 3730  
 3735  
 3740  
 3745  
 3750  
 3755  
 3760  
 3765  
 3770  
 3775  
 3780  
 3785  
 3790  
 3795  
 3800  
 3805  
 3810  
 3815  
 3820  
 3825  
 3830  
 3835  
 3840  
 3845  
 3850  
 3855  
 3860  
 3865  
 3870  
 3875  
 3880  
 3885  
 3890  
 3895  
 3900  
 3905  
 3910  
 3915  
 3920  
 3925  
 3930  
 3935  
 3940  
 3945  
 3950  
 3955  
 3960  
 3965  
 3970  
 3975  
 3980  
 3985  
 3990  
 3995  
 4000  
 4005  
 4010  
 4015  
 4020  
 4025  
 4030  
 4035  
 4040  
 4045  
 4050  
 4055  
 4060  
 4065  
 4070  
 4075  
 4080  
 4085  
 4090  
 4095  
 4100  
 4105  
 4110  
 4115  
 4120  
 4125  
 4130  
 4135  
 4140  
 4145  
 4150  
 4155  
 4160  
 4165  
 4170  
 4175  
 4180  
 4185  
 4190  
 4195  
 4200  
 4205  
 4210  
 4215  
 4220  
 4225  
 4230  
 4235  
 4240  
 4245  
 4250  
 4255  
 4260  
 4265  
 4270  
 4275  
 4280  
 4285  
 4290  
 4295  
 4300  
 4305  
 4310  
 4315  
 4320  
 4325  
 4330  
 4335  
 4340  
 4345  
 4350  
 4355  
 4360  
 4365  
 4370  
 4375  
 4380  
 4385  
 4390  
 4395  
 4400  
 4405  
 4410  
 4415  
 4420  
 4425  
 4430  
 4435  
 4440  
 4445  
 4450  
 4455  
 4460  
 4465  
 4470  
 4475  
 4480  
 4485  
 4490  
 4495  
 4500  
 4505  
 4510  
 4515  
 4520  
 4525  
 4530  
 4535  
 4540  
 4545  
 4550  
 4555  
 4560  
 4565  
 4570  
 4575  
 4580  
 4585  
 4590  
 4595  
 4600  
 4605  
 4610  
 4615  
 4620  
 4625  
 4630  
 4635  
 4640  
 4645  
 4650  
 4655  
 4660  
 4665  
 4670  
 4675  
 4680  
 4685  
 4690  
 4695  
 4700  
 4705  
 4710  
 4715  
 4720  
 4725  
 4730  
 4735  
 4740  
 4745  
 4750  
 4755  
 4760  
 4765  
 4770  
 4775  
 4780  
 4785  
 4790  
 4795  
 4800  
 4805  
 4810  
 4815  
 4820  
 4825  
 4830  
 4835  
 4840  
 4845  
 4850  
 4855  
 4860  
 4865  
 4870  
 4875  
 4880  
 4885  
 4890  
 4895  
 4900  
 4905  
 4910  
 4915  
 4920  
 4925  
 4930  
 4935  
 4940  
 4945  
 4950  
 4955  
 4960  
 4965  
 4970  
 4975  
 4980  
 4985  
 4990  
 4995  
 5000  
 5005  
 5010  
 5015  
 5020  
 5025  
 5030  
 5035  
 5040  
 5045  
 5050  
 5055  
 5060  
 5065  
 5070  
 5075  
 5080  
 5085  
 5090  
 5095  
 5100  
 5105  
 5110  
 5115  
 5120  
 5125  
 5130  
 5135  
 5140  
 5145  
 5150  
 5155  
 5160  
 5165  
 5170  
 5175  
 5180  
 5185  
 5190  
 5195  
 5200  
 5205  
 5210  
 5215  
 5220  
 5225  
 5230  
 5235  
 5240  
 5245  
 5250  
 5255  
 5260  
 5265  
 5270  
 5275  
 5280  
 5285  
 5290  
 5295  
 5300  
 5305  
 5310  
 5315  
 5320  
 5325  
 5330  
 5335  
 5340  
 5345  
 5350  
 5355  
 5360  
 5365  
 5370  
 5375  
 5380  
 5385  
 5390  
 5395  
 5400  
 5405  
 5410  
 5415  
 5420  
 5425  
 5430  
 5435  
 5440  
 5445  
 5450  
 5455  
 5460  
 5465  
 5470  
 5475  
 5480  
 5485  
 5490  
 5495  
 5500  
 5505  
 5510  
 5515  
 5520  
 5525  
 5530  
 5535  
 5540  
 5545  
 5550  
 5555  
 5560  
 5565  
 5570  
 5575  
 5580  
 5585  
 5590  
 5595  
 5600  
 5605  
 5610  
 5615  
 5620  
 5625  
 5630  
 5635  
 5640  
 5645  
 5650  
 5655  
 5660  
 5665  
 5670  
 5675  
 5680  
 5685  
 5690  
 5695  
 5700  
 5705  
 5710  
 5715  
 5720  
 5725  
 5730  
 5735  
 5740  
 5745  
 5750  
 5755  
 5760  
 5765  
 5770  
 5775  
 5780  
 5785  
 5790  
 5795  
 5800  
 5805  
 5810  
 5815  
 5820  
 5825  
 5830  
 5835  
 5840  
 5845  
 5850  
 5855  
 5860  
 5865  
 5870  
 5875  
 5880  
 5885  
 5890  
 5895  
 5900  
 5905  
 5910  
 5915  
 5920  
 5925  
 5930  
 5935  
 5940  
 5945  
 5950  
 5955  
 5960  
 5965  
 5970  
 5975  
 5980  
 5985  
 5990  
 5995  
 6000  
 6005  
 6010  
 6015  
 6020  
 6025  
 6030  
 6035  
 6040  
 6045  
 6050  
 6055  
 6060  
 6065  
 6070  
 6075  
 6080  
 6085  
 6090  
 6095  
 6100  
 6105  
 6110  
 6115  
 6120  
 6125  
 6130  
 6135  
 6140  
 6145  
 6150  
 6155  
 6160  
 6165  
 6170  
 6175  
 6180  
 6185  
 6190  
 6195  
 6200  
 6205  
 6210  
 6215  
 6220  
 6225  
 6230  
 6235  
 6240  
 6245  
 6250  
 6255  
 6260  
 6265  
 6270  
 6275  
 6280  
 6285  
 6290  
 6295  
 6300  
 6305  
 6310  
 6315  
 6320  
 6325  
 6330  
 6335  
 6340  
 6345  
 6350  
 6355  
 6360  
 6365  
 6370  
 6375  
 6380  
 6385  
 6390  
 6395  
 6400  
 640



347

Gln Ala Leu Leu Ser Asn Gly Phe Lys Lys Val Thr Phe Lys Pro Tyr  
115 120  
Ser Ala Leu Gly His Thr Ile Pro Gln Glu Leu Asp Glu Leu Cys  
130 135  
5 Ala Trp Leu Thr Ser Thr Leu Ser Leu Glu Gly  
140 145 150 155

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 643 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..643

(D) OTHER INFORMATION: / Ceres Seq. ID 1011631

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

agaatgtgat ggaatgga gtaataaga tgaatgtg gattaaga gagaagaag 60  
agaaggtct tatcttga ttgaactg atcagatc atcccttgg gaggccaa 120  
ggatcccatg gacgcgag ggcgcctc gtaataact cagatggt tgttcccaa 180  
ccatccat ggaagaag ggaagagc ccatcata caaccact cgcgcctag 240  
gttacact agagatgt ggaatgag gaagacagc taggtttca agatccgag 300  
agaaaagay gacaaggtg tttctcaga agataagga cgaagtact aaattgaatg 360  
cagagaagay gcttcgat agaaaggt cgcgaagay atcttcaat ggttgcctc 420  
actaagaac ttaataat atggataa aattacttg cttcatctt ctttttgt 480  
tctatagt ttgtgattg ttgattct ttttcgat tcatagaga ttgtgacgt 540  
actatttat atagtctga taactgaat ttgtattgt ggc 600

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 140 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..140

(D) OTHER INFORMATION: / Ceres Seq. ID 1011632

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Asn Val Met Glu Met Gly Val Asn Glu Met Ser Gly Gly Ile Lys Glu  
1 5 10 15  
Glu Lys Lys Glu Lys Ala Leu Met Leu Arg Leu Asp Tyr Glu Ser Val  
20 25 30  
Ile Ser Thr Trp Gly Gly Gln Gly Ile Pro Trp Thr Ala Arg Val Pro  
35 40 45  
Ser Glu Ile Asp Leu Asp Met Val Cys Phe Pro Thr His Thr Met Gly  
50 55 60  
Glu Ser Gly Ala Glu Ala His His Asn His Phe Arg Gly Leu Gly  
65 70 75  
Leu His Leu Gly Asp Ala Gly Asp Gly Arg Glu Ala Arg Val Ser  
80 85 90  
Arg Tyr Arg Glu Lys Arg Arg Thr Arg Leu Phe Ser Lys Lys Ile Arg  
95 100 105  
Tyr Glu Val Arg Lys Leu Asn Ala Glu Lys Arg Pro Arg Met Lys Gly  
110 115 120 125  
Arg Phe Val Lys Arg Ser Ile Gly Val Ala His  
130 135 140

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 138 amino acids

348

(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

5

(A) NAME/KEY: peptide  
(B) LOCATION: 1..138  
(D) OTHER INFORMATION: / Ceres Seq. ID 1011633  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Met Glu Met Gly Val Asn Glu Met Ser Gly Gly Ile Lys Glu Glu Lys  
1 5 10 15  
Lys Glu Lys Ala Leu Met Leu Arg Leu Asp Tyr Glu Ser Val Ile Ser  
20 25 30  
Thr Trp Gly Gly Gln Gly Ile Pro Trp Thr Ala Arg Val Pro Ser Glu  
35 40 45  
Ile Asp Leu Asp Met Val Cys Phe Pro Thr His Thr Met Gly Glu Ser  
50 55 60  
Gly Ala Glu Ala His His Asn His Phe Arg Gly Leu Gly Leu His  
65 70 75 80  
Leu Gly Asp Ala Gly Asp Gly Arg Glu Ala Arg Val Ser Arg Tyr  
90 95  
Arg Glu Lys Arg Arg Thr Arg Leu Phe Ser Lys Lys Ile Arg Tyr Glu  
100 105 110  
Val Arg Lys Leu Asn Ala Glu Lys Arg Pro Arg Met Lys Gly Arg Phe  
115 120 125  
Val Lys Arg Ser Ile Gly Val Ala His  
130 135

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 136 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..136

(D) OTHER INFORMATION: / Ceres Seq. ID 1011634

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

Met Gly Val Asn Glu Met Ser Gly Ile Lys Glu Glu Lys Lys Glu  
1 5 10 15  
Lys Ala Leu Met Leu Arg Leu Asp Tyr Glu Ser Val Ile Ser Thr Trp  
20 25 30  
Gly Gly Gln Gly Ile Pro Trp Thr Ala Arg Val Pro Ser Glu Ile Asp  
35 40 45  
Leu Asp Met Val Cys Phe Pro Thr His Thr Met Gly Glu Ser Gly Ala  
50 55 60  
Glu Ala His His Asn His Phe Arg G-y Leu Gly Leu His Leu Gly  
65 70 75 80  
Asp Ala Gly Asp Gly Arg Glu Ala Arg Val Ser Arg Tyr Arg Glu  
90 95  
Lys Arg Arg Thr Arg Leu Phe Ser Lys Lys Ile Arg Tyr Glu Val Arg  
100 105 110  
Lys Leu Asn Ala Glu Lys Arg Pro Arg Met Lys Gly Arg Phe Val Lys  
115 120 125  
Arg Ser Ser Ile Gly Val Ala His  
130 135

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 663 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

349

(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

5

(A) NAME/KEY: -  
(B) LOCATION: 1..663  
(D) OTHER INFORMATION: / Ceres Seq. ID 1011714  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

gtgttcgata ataaagtcga cagagaagc gatacaaga gtcgcgtgctg tctgtttcat 60  
tctcaacgga aactcaatct tatcagtgct cscgcgttcc taaatcgga actccacttt 120  
cgctcttccc gttggccacc tgaatctcgg agagagcttc gaagaatgtg cagcagagga 180  
agtaacggag gaacagcgtc taagaatgga aaagatgaag cttttgactg ttacaacaa 240  
gtgttcaaa gaagcaacaa cgcacacaca ctacgtctct gttctgatac gtcgcgtgtt 300  
gtatgattgg gagactatcc cgaagaatat ggaacagag aagtggtgaag gtcggatttg 360  
tggttcaat ccttctatcc caaagccttt gtttggcca cttgagaaat ttttgggaag 420  
tttggatttg aatgttccaa atgtggcatt tggttatgtg atgtcaaggt taatattga 480  
gcagaggttc cgaagtcca ttctcggaat gccctctct tttactcttt ttattgtat 540  
tttggtttc tttttctcta attagtaata aacttatcat gtgtcctatc gattttattc 660  
ctc

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 151 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:

(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..151

(D) OTHER INFORMATION: / Ceres Seq. ID 1011715  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

Cys Ser Ile Ile Met Ser Thr Gly Glu Ala Ile Pro Arg Val Ala Val 15  
Val Val Phe Ile Leu Asn Gly Asn Ser Ile Leu Leu Gly Arg Arg Arg 35  
Ser Ser Ile Gly Asn Ser Thr Phe Ala Leu Pro Gly Gly His Leu Glu 45  
Phe Gly Glu Ser Phe Glu Glu Cys Ala Ala Arg Glu Val Met Glu Glu 50  
Thr Gly Leu Lys Ile Glu Lys Met Lys Leu Leu Thr Val Thr Asn Asn 60  
Val Phe Lys Glu Ala Pro Thr Pro Ser His Tyr Val Ser Val Ser Ile 75  
Arg Ala Val Leu Val Asp Pro Ser Glu Glu Pro Lys Asn Met Glu Pro 85  
Glu Lys Cys Glu Gly Trp Asp Trp Tyr Asp Trp Glu Asn Leu Pro Lys 100  
Pro Leu Phe Trp Pro Leu Glu Lys Leu Phe Gly Ser Gly Phe Asn Pro 110  
Phe Thr His Gly Gly Asp 120  
145

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 147 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:

(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..147

(D) OTHER INFORMATION: / Ceres Seq. ID 1011716

350

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

Met Ser Thr Gly Glu Ala Ile Pro Arg Val Ala Val Val Phe Ile 15  
Leu Asn Gly Asn Ser Ile Leu Leu Gly Arg Arg Ser Ile Gly 30  
Asn Ser Thr Phe Ala Leu Pro Gly Gly His Leu Glu Phe Gly Glu Ser 45  
Phe Glu Glu Cys Ala Ala Arg Glu Val Met Glu Glu Thr Gly Leu Lys 60  
Ile Glu Lys Met Lys Leu Thr Val Thr Asn Asn Val Phe Lys Glu 75  
Ala Pro Thr Pro Ser His Tyr Val Ser Val Ser Ile Arg Ala Val Leu 90  
Val Asp Pro Ser Glu Glu Pro Lys Asn Met Glu Pro Glu Lys Cys Glu 105  
Gly Trp Asp Trp Tyr Asp Trp Glu Asn Leu Pro Lys Pro Lys Phe Trp 120  
Pro Leu Glu Lys Leu Phe Gly Ser Gly Phe Asn Pro Phe Thr His Gly 135  
Gly Gly Asp 140  
145

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 90 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:

(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..90

(D) OTHER INFORMATION: / Ceres Seq. ID 1011717

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

Met Glu Glu Thr Gly Leu Lys Ile Glu Lys Met Lys Leu Leu Thr Val 15  
Thr Asn Asn Val Phe Lys Glu Ala Pro Thr Pro Ser His Tyr Val Ser 30  
Val Ser Ile Arg Ala Val Leu Val Asp Pro Ser Glu Glu Pro Lys Asn 45  
Met Glu Pro Glu Lys Cys Glu Gly Trp Asp Trp Tyr Asp Trp Glu Asn 60  
Leu Pro Lys Pro Leu Phe Thr Pro Leu Cys Lys Leu Phe Gly Ser Gly 75  
Phe Asn Pro Phe Thr His Gly Gly Asp 90  
85

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 524 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..524

(D) OTHER INFORMATION: / Ceres Seq. ID 1011784

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

aaacagaac aaacacaac acaacacag aatcacaga gaagaaaat ggtgaagtg 60  
agtgggttt cgtttttagc tctggcggcg cgtatctccc tttagcgggt ccggtggca 120  
gaagggtga cgtctccacc tatcagctg gcttcatgt cggcggcgat gaagtcact 180  
tcgcgccat cggagcgtg ttgcacaag ctgagcagc agcagcgtg cctttgtgg 240  
tcacatgaga accttacct ccgtcaatac cttagctccc cttaacgag gaagtcctc 300

351

aaagttgca agattcttc cccagctgt taaggaaatg taatcatga ttaattagtg 360  
 accagtttgc cgtattatag tggtaagtc tggcttatc ttgcagta ctactaaa 420  
 taataatga cgtatgata ttgatgata tcatcatat gctgaatg: tatgttttc 480  
 acctataaa acatgaata tctttgaa cagtttgc tegt

(2) INFORMATION FOR SEQ ID NO:55:  
 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 110 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..110

(D) OTHER INFORMATION: / Ceres Seq. ID 1011785

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

1 Lys Thr Glu Gln Thr Gln Thr Gln Thr Glu Asn Gln Arg Arg Arg Lys 15  
 5 Met Val Lys Val Met Trp Val Ser Val Leu Ala Leu Ala Ala Ile 20  
 10 Leu Leu Thr Val Pro Val Ala Glu Gly Val Thr Cys Ser Pro Met: 25  
 35 Gln Leu Ala Ser Cys Ala Ala Met Thr Ser Ser Ser Pro Ser 40  
 50 Glu Ala Cys Cys Thr Lys Leu Arg Glu Gln Gln Pro Cys Leu Cys Gly 45  
 65 Tyr Met Arg Asn Pro Thr Leu Arg Gln Tyr Val Ser Pro Asn Ala 50  
 85 Arg Lys Val Ser Asn Ser Cys Lys Ile Pro Ser Pro Ser Cys 105  
 110

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 94 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..94

(D) OTHER INFORMATION: / Ceres Seq. ID 1011706

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

1 Met Val Lys Val Met Trp Val Ser Val Leu Ala Leu Ala Ala Ala Ile 15  
 5 Leu Leu Thr Val Pro Val Ala Glu Gly Val Thr Cys Ser Pro Met: 20  
 25 Gln Leu Ala Ser Cys Ala Ala Met Thr Ser Ser Ser Pro Ser 30  
 35 Glu Ala Cys Cys Thr Lys Leu Arg Glu Gln Gln Pro Cys Leu Cys Gly 40  
 45 Tyr Met Arg Asn Pro Thr Leu Arg Gln Tyr Val Ser Pro Asn Ala 45  
 65 Arg Lys Val Ser Asn Ser Cys Lys Ile Pro Ser Pro Ser Cys 55  
 85

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 90 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..90

(D) OTHER INFORMATION: / Ceres Seq. ID 1011821

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

1 Pro Asn Phe Ser Met Asp Arg Ser Leu Leu Leu Leu Leu Leu Leu 15  
 5 Gly Phe Asp Glu Gln Leu Ala Ile Pro Ser Pro Leu Asn Gly Lys Val: 20  
 25 Ile Leu Ile Glu Asp Cys Val Glu Thr Ser Gly Ser Phe Val Leu His 30  
 35

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 107 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..1007

(D) OTHER INFORMATION: / Ceres Seq. ID 1011820

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

1 ccgaattctt ccatggaccg tctttgaat ctctcgatt tagcttcagg gttcgtgag 60  
 120 cagctgcta tccatctgc actaatgga aaagaaatc taatagaaga ctgtgtgag 180  
 240 agagtggtt ctttctact tcaacagta atgaacgtg tctctctc taactctcc 240  
 300 gacgactta tcttctgc tttgtctgc cttctctc ctctctctc attatgctg aattctggt 360  
 420 aaactggat gtaatttagc taccataag tcaataatc gattggtt ctttgatag 360  
 480 cctatgta agcttcaga tgggtacaa atggaagaa atgagagc agttgcgaa 420  
 540 actatcgg agatacaaga aaccttga aagctacaga gtgtaacag tggtaacata 480  
 600 caagatgg acttttga tttctctg gaattgcta ctacggcg caactagat 540  
 660 ctcatctca atcatgaaga tatacagg agcatggga gacttcatt ttgtctag 600  
 660 aggtatgcc ttgcagatg ttgtataag gcagagcct tagcttcag tttagcaat 660  
 720 gatgacatg gccaatgac tttctgac aaaggataa gcaactcag tagagaagc 720  
 780 tggagaaca agttgcaga ttctcaattc aggtcaag aaatggtat cgtattctc 780  
 840 tatctggtt gcagaagtg agatagac acgttctga gattagcat ggtatgaga 840  
 900 tgcctaga cagtttgtt attgctgtt aggttaaca gagacactt gttcaacat 900  
 960 ttctcttag tctctctt ggttgaat tgatgacta gctaaagtg aatgcttag 960  
 acacagaac ttgagacat ggaattag taacattcta tatcttt

352

(A) NAME/KEY: peptide  
 (B) LOCATION: 1..90  
 (C) OTHER INFORMATION: / Ceres Seq. ID 1011787

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

1 Asn Arg Thr Asn Thr Asn Thr Asn Arg Glu Ser Thr Lys Lys Lys Asn 15  
 5 Gly Glu Gly Asp Val Gly Phe Arg Phe Ser Ser Gly Gly Gly Asp Ser 20  
 25 Pro Phe Asp Gly Pro Gly Gly Arg Arg Gly Asp Val Leu Ala Tyr Ala 30  
 35 Ala Gly Phe Met Cys Gly Gly Asp Asp Val Ile Phe Ala Ala Ile Gly 40  
 45 Gly Val Leu His Lys Lys Ala Glu Arg Ala Ala Met Pro Leu Trp Val 50  
 55 His Glu Glu Pro Tyr Pro Pro Ile Arg 60  
 65 70 75 80 85 90

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1007 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..1007

(D) OTHER INFORMATION: / Ceres Seq. ID 1011820

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

1 ccgaattctt ccatggaccg tctttgaat ctctcgatt tagcttcagg gttcgtgag 60  
 120 cagctgcta tccatctgc actaatgga aaagaaatc taatagaaga ctgtgtgag 180  
 240 agagtggtt ctttctact tcaacagta atgaacgtg tctctctc taactctcc 240  
 300 gacgactta tcttctgc tttgtctgc cttctctc ctctctctc attatgctg aattctggt 360  
 420 aaactggat gtaatttagc taccataag tcaataatc gattggtt ctttgatag 360  
 480 cctatgta agcttcaga tgggtacaa atggaagaa atgagagc agttgcgaa 420  
 540 actatcgg agatacaaga aaccttga aagctacaga gtgtaacag tggtaacata 480  
 600 caagatgg acttttga tttctctg gaattgcta ctacggcg caactagat 540  
 660 ctcatctca atcatgaaga tatacagg agcatggga gacttcatt ttgtctag 600  
 660 aggtatgcc ttgcagatg ttgtataag gcagagcct tagcttcag tttagcaat 660  
 720 gatgacatg gccaatgac tttctgac aaaggataa gcaactcag tagagaagc 720  
 780 tggagaaca agttgcaga ttctcaattc aggtcaag aaatggtat cgtattctc 780  
 840 tatctggtt gcagaagtg agatagac acgttctga gattagcat ggtatgaga 840  
 900 tgcctaga cagtttgtt attgctgtt aggttaaca gagacactt gttcaacat 900  
 960 ttctcttag tctctctt ggttgaat tgatgacta gctaaagtg aatgcttag 960  
 acacagaac ttgagacat ggaattag taacattcta tatcttt

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 266 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..266

(D) OTHER INFORMATION: / Ceres Seq. ID 1011821

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

1 Pro Asn Phe Ser Met Asp Arg Ser Leu Leu Leu Leu Leu Leu Leu 15  
 5 Gly Phe Asp Glu Gln Leu Ala Ile Pro Ser Pro Leu Asn Gly Lys Val: 20  
 25 Ile Leu Ile Glu Asp Cys Val Glu Thr Ser Gly Ser Phe Val Leu His 30  
 35

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 266 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..266

(D) OTHER INFORMATION: / Ceres Seq. ID 1011821

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

1 Pro Asn Phe Ser Met Asp Arg Ser Leu Leu Leu Leu Leu Leu Leu 15  
 5 Gly Phe Asp Glu Gln Leu Ala Ile Pro Ser Pro Leu Asn Gly Lys Val: 20  
 25 Ile Leu Ile Glu Asp Cys Val Glu Thr Ser Gly Ser Phe Val Leu His 30  
 35

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 266 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..266

(D) OTHER INFORMATION: / Ceres Seq. ID 1011821

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

1 Pro Asn Phe Ser Met Asp Arg Ser Leu Leu Leu Leu Leu Leu Leu 15  
 5 Gly Phe Asp Glu Gln Leu Ala Ile Pro Ser Pro Leu Asn Gly Lys Val: 20  
 25 Ile Leu Ile Glu Asp Cys Val Glu Thr Ser Gly Ser Phe Val Leu His 30  
 35

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 266 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..266

(D) OTHER INFORMATION: / Ceres Seq. ID 1011821

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

1 Pro Asn Phe Ser Met Asp Arg Ser Leu Leu Leu Leu Leu Leu Leu 15  
 5 Gly Phe Asp Glu Gln Leu Ala Ile Pro Ser Pro Leu Asn Gly Lys Val: 20  
 25 Ile Leu Ile Glu Asp Cys Val Glu Thr Ser Gly Ser Phe Val Leu His 30  
 35

35 353 45  
Gln Leu Met Lys Arg Val Leu Ser Ser Asn Ser Ser Asp Ala Leu Ile  
50 55 60  
Phe Leu Ala Phe Ala Arg Pro Phe Ser His Tyr Asp Arg Ile Leu Arg  
65 70 75 80  
Lys Leu Gly Cys Asn Leu Ala Thr His Lys Ser Asn Asn Arg Leu Val  
85 90 95  
Phe Phe Asp Met Leu Met Val Lys Cys Ser Asp Gly Asp Gln Met Glu  
100 105 110  
10 Asp Asn Val Ser Ala Val Ala Lys Leu Phe Arg Glu Ile Gln Glu Thr  
115 120 125  
Val Arg Lys Leu Gln Ser Val Thr Ser Gly Asn Ile Thr Val Met Val  
130 135 140  
15 Asp Asp Met Ser Leu Leu Glu Ile Ala Thr Thr Gly Ser Asn Ser Asp  
145 150 155 160  
His Val Leu Asp Phe Leu His Tyr Cys His Thr Leu Ser Ser Glu Ser  
165 170 175  
Asn Cys Ser Leu Val Ile Leu Asn His Glu Asp Ile Tyr Ala Ser Met  
180 185 190  
20 Glu Arg Pro Ala Phe Leu Leu Gln Met Val Cys Leu Ala Asp Val Val  
195 200 205  
Ile Lys Ala Glu Pro Leu Ala Ser Gly Leu Ala Asn Asp Val His Gly  
210 215 220  
Gln Leu Thr Val Leu Asn Lys Gly Ile Ser Asn Ser Gly Arg Gly Ser  
225 230 235 240  
Ser Arg Asn Lys Leu Gln Asn Phe Gln Phe Arg Ile Lys Glu Asn Gly  
245 250 255  
Ile Asp Tyr Phe Tyr Pro Gly Cys Arg Ser  
260 265

(2) INFORMATION FOR SEQ ID NO:60:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 262 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(11) MOLECULE TYPE: peptide  
(1x) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..262  
(D) OTHER INFORMATION: / Ceres Seq. ID 1011822  
(xl) SEQUENCE DESCRIPTION: SEQ ID NO:60:  
Met Asp Arg Ser Leu Asn Leu Ser Asp Leu Ala Leu Gly Phe Asp Glu  
1 5 10 15  
Gln Leu Ala Ile Pro Ser Pro Leu Asn Gly Lys Val Ile Leu Ile Glu  
20 25 30  
Asp Cys Val Glu Thr Ser Gly Ser Phe Val Leu His Gln Leu Met Lys  
35 40 45  
Arg Val Leu Ser Ser Asn Ser Asp Ala Leu Ile Phe Leu Ala Phe  
50 55 60  
Ala Arg Pro Phe Ser His Tyr Asp Arg Ile Leu Arg Lys Leu Gly Cys  
65 70 75 80  
Asn Leu Ala Thr His Lys Ser Asn Asn Arg Leu Val Phe Phe Asp Met  
85 90 95  
Leu Met Val Lys Cys Ser Asp Gly Asp Gln Met Glu Asp Asn Val Ser  
100 105 110  
Ala Val Ala Lys Leu Phe Arg Glu Ile Gln Glu Thr Val Arg Lys Leu  
115 120 125  
Gln Ser Val Thr Ser Gly Asn Ile Thr Val Met Val Asp Asp Met Ser  
130 135 140  
Leu Leu Glu Ile Ala Thr Thr Gly Ser Asn Ser Asp His Val Leu Asp  
145 150 155 160  
Phe Leu His Tyr Cys His Thr Leu Ser Ser Glu Ser Asn Cys Ser Leu

165 170 175 354  
Val Ile Leu Asn His Glu Asp Ile Tyr Ala Ser Met Glu Arg Pro Ala  
180 185 190  
Phe Leu Leu Gln Met Val Cys Leu Ala Asp Val Val Ile Lys Ala Glu  
195 200 205  
Pro Leu Ala Ser Gly Leu Ala Asn Asp Val His Gly Gln Leu Thr Val  
210 215 220  
Leu Asn Lys Gly Ile Ser Asn Ser Gly Arg Gly Ser Ser Arg Asn Lys  
225 230 235 240  
Leu Gln Asn Phe Gln Phe Arg Ile Lys Glu Asn Gly Ile Asp Tyr Phe  
245 250 255  
Tyr Pro Gly Cys Arg Ser  
260 265

(2) INFORMATION FOR SEQ ID NO:61:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 216 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(11) MOLECULE TYPE: peptide  
(1x) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..216  
(D) OTHER INFORMATION: / Ceres Seq. ID 1011023  
(xl) SEQUENCE DESCRIPTION: SEQ ID NO:61:  
Met Lys Arg Val Leu Ser Ser Asn Ser Asp Ala Leu Ile Phe Leu  
1 5 10 15  
Ala Phe Ala Arg Pro Phe Ser His Tyr Asp Arg Ile Leu Arg Lys Leu  
20 25 30  
Gly Cys Asn Leu Ala Thr His Lys Ser Asn Asn Arg Leu Val Phe Phe  
35 40 45  
Asp Met Leu Met Val Lys Cys Ser Asp Gly Asp Gln Met Glu Asp Asn  
50 55 60  
Val Ser Ala Val Ala Lys Leu Phe Arg Glu Ile Gln Glu Thr Val Arg  
65 70 75 80  
Lys Leu Gln Ser Val Thr Ser Gly Asn Ile Thr Val Met Val Asp Asp  
85 90 95  
Met Ser Leu Leu Glu Ile Ala Thr Thr Gly Ser Asn Ser Asp His Val  
100 105 110  
Leu Asp Phe Leu His Tyr Cys His Thr Leu Ser Ser Glu Ser Asn Cys  
115 120 125  
Ser Leu Val Ile Leu Asn His Glu Asp Ile Tyr Ala Ser Met Glu Arg  
130 135 140  
Pro Ala Phe Leu Leu Gln Met Val Cys Leu Ala Asp Val Val Ile Lys  
145 150 155 160  
Ala Glu Pro Leu Ala Ser Gly Leu Ala Asn Asp Val His Gly Gln Leu  
170 175  
Thr Val Leu Asn Lys Gly Ile Ser Asn Ser Gly Arg Gly Ser Ser Arg  
180 185 190  
Asn Lys Leu Gln Asn Phe Gln Phe Arg Ile Lys Glu Asn Gly Ile Asp  
195 200 205  
Tyr Phe Tyr Pro Gly Cys Arg Ser  
210 215 220

(2) INFORMATION FOR SEQ ID NO:62:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 584 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(11) MOLECULE TYPE: DNA (genomic)  
(1x) FEATURE:  
(A) NAME/KEY: -

355

(B) LOCATION: 1..584  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011874  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:  
 5 tcaattaca taactttgc cttgatgaa gctttccct caattccac taccttcacg 60  
 gtcgaacga acaacaacta cagcgaataa ccaataccac atgaagttta tgaaggagaa 120  
 aggaagatag acgacagag caccagtagt acgcccaggt tgcgggttg ggaagtcca 180  
 gaagactag ggaacttct tagacttgc cttgatgaa agaatggac cgaatgcaat 240  
 attcgaaga tgaactatc tgttcaatc ggaactcag tgaatgacc aggcatttc 300  
 catgacatt atatagtc tggtttga atgattcac ttgatctca aatccattac 360  
 gtaaacatt ttggaact ttittttt ttittttta ttggtatg ttccaagaa 420  
 aaatggtt ctggaatt aattatag tccgtatg ttggtatg ttccaagaa 480  
 taagaact taagttgt tagagttt taacattga aattttaca aaattctcat 540  
 taaagtgt taacaact ttgtgcat gatttggtg ttgt

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 135 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..135

(D) OTHER INFORMATION: / Ceres Seq. ID 1011875

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

1 Ser Ile Asn Ile Thr Phe Ala Val Asp Glu Ala Phe Pro Ser Ile Pro 15  
 5 Thr Thr Phe Ser Val Ala Thr Lys Gln His Tyr Asp Val Lys Pro Ile 20  
 20 His His Glu Val Tyr Asp Gly Glu Arg Lys Ile Tyr Asp Ile Ser His 35  
 35 Gln Tyr Thr Pro Glu Leu Pro Val Trp Glu Ser Ser Glu Gly Leu Gly 40  
 40 Asn Phe Leu Arg Leu Ala Val Ser Met Lys Asn Gly Ser Asp Ala Asn 50  
 50 Ile Ser Lys Met Glu Leu Ser Val His Ser Gly Thr His Val Asp Ala 60  
 60 Pro Gly His Phe His Asp His Tyr Tyr Glu Ser Gly Phe Asp Thr Asp 75  
 75 Ser Leu Asp Leu Gln Ile Leu Asn Gly Lys His Phe Thr Asn Ser Phe 85  
 85 Phe Phe Phe Leu Ser Ser 120  
 120 Phe Phe Phe Leu Ser Ser 135

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 626 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..626

(D) OTHER INFORMATION: / Ceres Seq. ID 1011981

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

55 aatcttaat cgaataaac cctagaanaa acctctctt cctctcaat tcaattctc 60  
 tctacaag gtttgga acgagagac tgcgtctgt tcttgctt cgcacacg 120  
 aagatacca tctctcgt cctccacac cgcctcgt cgtctcgt cctctaatc 180  
 ctcttaact agcgtggt tcaattct gtaactcc aggaattag gacccctag 240  
 ttgcacag tggttctgc ccttgtag tgttggtt agttctac tgacattca 300  
 cctcaagct aattggag cttctcga gttctaac gttctaac ggtctgaa gaggcaag 360  
 agtg-aatt cactgcgat gcgtgaatca gtaattgaa ggaatgtag atcgaaattg 420

356

tggaattacc tcaagaaaa cttgcttaat ggtgtcctg tttagaagt tttagttcca 480  
 caagtgtg attattagt tgaagacct tctatgctt tgcacaatg gttttctc 540  
 agagaacct tttttttt ggagaatgc acagcagat taataagtg attggtacg 600  
 aactgaaca atgtttac ttttc

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 93 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..93

(D) OTHER INFORMATION: / Ceres Seq. ID 1011982

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

1 Met Ala Trp Arg Asn Ala Gly Ser Ala Ala Arg Ser Phe Val Ser Ala 15  
 5 Thr Ala Arg Ser Pro Ser Leu Arg Ser Pro Thr Thr Ala Leu Pro Arg 20  
 20 Leu Arg Pro Gln Ser Ser Leu Pro Ser Arg Arg Phe Thr Phe Ser 25  
 25 Ser Pro Ser Arg Asn Leu Gly Ala Leu Gly Cys Thr Gln Ser Phe Leu 30  
 30 Pro Leu Tyr Ser Val Val Ala Thr Ser Gln Leu Thr Ser His Leu Asn 35  
 35 Val Asn Leu Arg Ala Phe Cys Glu Leu Ser Asn Gly Thr 40  
 40

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 75 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..75

(D) OTHER INFORMATION: / Ceres Seq. ID 1011983

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

1 Ile Leu Ile Glu Lys Asn Pro Ser Lys Asn Leu Ser Leu Ser Leu Asn 15  
 5 Ser Leu Leu Leu Leu Gln Trp Leu Gly Ala Thr Gln Asp Leu Leu 20  
 20 Val Leu Ser Ser Pro Pro Gln Asp His His Leu Ser Val Leu Pro 25  
 25 Pro Arg Arg Phe Leu Ala Ser Val Leu Leu Asn Pro Pro Tyr Leu Ala 30  
 30 Val Ala Ser Pro Ser Arg His Leu Pro Gly Ile 35  
 35

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 46 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..46

(D) OTHER INFORMATION: / Ceres Seq. ID 1011984

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

357

Met Cys Gly Ile Phe Ser Cys Arg Ala Phe Arg Met Leu Gly Thr Met  
 1 5 10 15  
 Val Phe Ser Gln Arg Thr Phe Phe Cys Gly Glu Met His Ser Ser  
 20 25 30  
 5 Ile Asn Lys Cys Ile Gly Ser Glu Leu Lys Gln Cys Cys Tyr  
 35 40 45

(2) INFORMATION FOR SEQ ID NO:68:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 48 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..468  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1014547

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

atagaatata accattatcc ggacatgaag atcgagcat tggtaactcgt cgtcttcgtc 60  
 atactttcga catcattccc gctgcgacc aaagccgaag acacgggaga tacaggaat 120  
 stggagatga catcgaccg aagcgactc cagctcgcc ttgcgcgat tagcgaaga 180  
 ggacacact cgggtagcgt ttgtgcaag cttaacgaqg acaactcatg ctatcggt 240  
 ttctgtaaga accc-gcgt tccacagttc attagcttc caaacgtog caaagctc 300  
 ctcttgata atgt-gctta tccactttg tgaacttat ctagatttta taataaata 360  
 aacgaaga aataattac attatctaa acgttatgt acaattcaac cgtttgtgtt 420  
 aatctatgt ctacatggt taataaagt ttaattctt ggttcgtc

(2) INFORMATION FOR SEQ ID NO:69:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 110 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..110  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1014548

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

Ile Glu Tyr Asn His Ile Pro Asp Met Cys Ile Val Thr Leu Val Leu  
 1 5 10 15  
 Val Val Phe Val Ile Leu Ser Thr Ser Phe Pro Ala Ala Ile Lys Ala  
 20 25 30  
 Glu Asp Thr Gly Asp Thr Gly Asn Val Gly Val Thr Cys Asp Ala Arg  
 35 40 45  
 Gln Leu Gln Pro Cys Leu Ala Ala Ile Thr Gly Gly Gln Pro Ser  
 50 55  
 Gly Ala Cys Cys Ala Lys Leu Thr Glu Gln Gln Ser Cys Leu Cys Gly  
 65 70 75 80  
 Phe Ala Lys Asn Pro Ala Phe Ala Gln Tyr Ile Ser Ser Pro Asn Ala  
 85 90 95  
 Arg Lys Val Leu Leu Ala Cys Asn Val Ala Tyr Pro Thr Cys  
 100 105 110

(2) INFORMATION FOR SEQ ID NO:70:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 102 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..102

358

(D) OTHER INFORMATION: / Ceres Seq. ID 1014549  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:  
 Met Lys Ile Val Thr Leu Val Leu Val Phe Val Ile Leu Ser Thr  
 1 5 10 15  
 Ser Phe Pro Ala Ala Ile Lys Ala Glu Asp Thr Cys Asp Thr Gly Asn  
 20 25 30  
 Val Gly Val Thr Cys Asp Ala Arg Gln Leu Gln Pro Cys Leu Ala Ala  
 35 40 45  
 Ile Thr Gly Gly Gly Gln Pro Ser Gly Ala Cys Cys Ala Lys Leu Thr  
 50 55 60  
 Glu Gln Gln Ser Cys Leu Cys Gly Phe Ala Lys Asn Pro Ala Phe Ala  
 65 70 75 80  
 Gln Tyr Ile Ser Ser Pro Asn Ala Arg Lys Val Leu Leu Ala Cys Asn  
 85 90 95  
 Val Ala Tyr Pro Thr Cys 100

(2) INFORMATION FOR SEQ ID NO:71:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1303 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..1303  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1014995

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

aaactcaag ttctttgcaa ttactactt acacaaagc aactgactc gaacaaacac 60  
 atcattgct cctctgcgc aactcttcc ttgtagtgt tccgatgaa aatctgcga 120  
 tgcacagg ggagagttg cattcaagt ttgttcagc gattatagt atgttgcga 180  
 gtgcaatcat ggagagttg ttcaagagg ggagattct ctttttgtt atgtgcact 240  
 taaccttgc gctgtgtc ttcaagagg ccaggttta tatgaaggac tgaagattc 300  
 caggacagaa gatgtcga ttctctatt ccgaccagac caaaaggc tcgccttca 360  
 agcggagct gacagactt atatgctta ttatcgttc gactcaatcg tctcgctat 420  
 caaacagtt gtctttgca aagaaalg gattctctt ccgggaaag gaacttgta 480  
 ttataggct atcttttg ggagtggtcc gattcttgt tcaillccca ttcttgagc 540  
 cactttaca gttttgct gtctgttgg agttatcat aagataact ctggtttgaa 600  
 tctgaaatc gaatcact ttctcgagc ttctctaat ggaactgtg gtytgaagag 660  
 catcaaac tatgtctcg ttgtatccc atggcgag ggcaaaac aggtttctc 720  
 tgatatatt tttttgatg ctgaactgg caaaacatt gaagacttt tgcagctaa 780  
 cggatcact cgaactcgg taatgaatt gttgtgat ttccgtacc aggttcgag 840  
 tgttttag ctcaaggca atgttgatc gacacaca attgcagaa ctattttgc 900  
 acgtcagat cctctagtg actttcga ttggagcaa gctttcgt ctggcagtc 960  
 ttcttttgt actgtattg catctaac ctttaagac aaaaagccg gattcaaac 1020  
 aggaagaaa cattgctg gaagctaac gtagcttaa gatatoca gacgggtcg 1080  
 gtcgagata ccaaggatg gacgttgag attgcgcc aggtcgaaa gttgaactg 1140  
 taacttgat ctaaatatg gtgtgtgtg atcaaaact tatgtaaga acatctgaag 1200  
 atgtctcga tcttttgtt tggatgac atcaatctc tatcgttgc argtatatta 1260  
 taactctcg tctgtaagt attataac tatgctttt tgc

(2) INFORMATION FOR SEQ ID NO:72:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 352 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..352  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1014996

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

359  
Lys Leu Lys Phe Leu Ala Ile Thr Tyr Leu Gln Gln Ser Asn Ser Thr  
1 5 10 15  
Arg Thr Asn Thr Ser Met Ala Pro Ser Ala Gln Pro Leu Pro Val Ser  
20 25 30  
5 Val Ser Asp Glu Lys Tyr Ala Asn Val Lys Trp Glu Glu Leu Ala Phe  
35 40 45  
Lys Phe Val Arg Thr Asp Tyr Met Tyr Val Ala Lys Cys Asn His Gly  
50 55 60  
Glu Ser Phe Gln Glu Gly Lys Ile Leu Pro Phe Ala Asp Leu Gln Leu  
65 70 75 80  
10 Asn Pro Cys Ala Ala Val Leu Gln Tyr Gly Gln Gly Leu Tyr Glu Gly  
85 90 95  
Leu Lys Ala Tyr Arg Thr Glu Asp Gly Arg Ile Leu Leu Phe Arg Pro  
100 105 110  
15 Asp Gln Asn Gly Leu Arg Leu Gln Ala Gly Ala Asp Arg Leu Tyr Met  
115 120 125  
Pro Tyr Pro Ser Val Asp Gln Phe Val Ser Ala Ile Lys Gln Val Ala  
130 135 140  
Leu Ala Asn Lys Lys Trp Ile Pro Pro Gly Lys Gly Thr Leu Tyr  
145 150 155  
Ile Arg Pro Ile Leu Phe Gly Ser Gly Pro Ile Leu Gly Ser Phe Pro  
160 165 170  
175  
Ile Pro Gln Thr Thr Phe Thr Ala Phe Ala Cys Pro Val Gly Arg Tyr  
180 185 190  
25 His Lys Asp Asn Ser Gly Leu Asn Leu Lys Ile Glu Asp Gln Phe Arg  
195 200 205  
Arg Ala Phe Pro Ser Gly Thr Gly Gly Val Lys Ser Ile Thr Asn Tyr  
210 215 220  
Cys Pro Val Trp Ile Pro Leu Ala Glu Ala Lys Lys Gln Gly Phe Ser  
225 230 235  
30 Asp Ile Leu Phe Leu Asp Ala Ala Thr Gly Lys Asn Ile Glu Glu Leu  
240 245 250  
Phe Ala Ala Asn Val Phe Met Leu Lys Gly Asn Val Val Ser Thr Pro  
255 260 265  
35 Thr Ile Ala Gly Thr Ile Leu Pro Gly Val Thr Arg Asn Cys Val Met  
270 275 280  
Glu Leu Cys Arg Asp Phe Gly Tyr Gln Val Glu Glu Arg Thr Ile Pro  
285 290 300  
Leu Val Asp Phe Leu Asp Ala Asp Glu Ala Phe Cys Thr Gly Thr Ala  
305 310 315  
40 Ser Ile Val Thr Ser Ile Ala Ser Val Thr Phe Lys Asp Lys Lys Thr  
320 325 330  
Gly Phe Lys Thr Gly Lys Lys His Trp Leu Arg Ser Tyr Thr Arg Arg  
335 340 345  
45 (2) INFORMATION FOR SEQ ID NO:73:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 331 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) FEATURE:  
(ii) MOLECULE TYPE: peptide  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..331  
(D) OTHER INFORMATION: / Ceres Seq. ID 1014997  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:  
Met Ala Pro Ser Ala Gln Pro Leu Pro Val Ser Val Ser Asp Glu Lys  
1 5 10  
Tyr Ala Asn Val Lys Trp Glu Glu Leu Ala Phe Lys Phe Val Arg Thr  
15 20 25  
Asp Tyr Met Tyr Val Ala Lys Cys Asn His Gly Glu Ser Phe Gln Glu  
30 35 40 45

360  
Gly Lys Ile Leu Pro Phe Ala Asp Leu Gln Leu Asn Pro Cys Ala Ala  
50 55 60  
Val Leu Gln Tyr Gly Gln Gly Leu Tyr Cys Gly Leu Lys Ala Tyr Arg  
65 70 75 80  
5 Thr Glu Asp Gly Ile Leu Leu Phe Arg Pro Asp Gln Asn Gly Leu  
85 90 95  
Arg Leu Gln Ala Gly Ala Asp Arg Leu Tyr Met Pro Tyr Pro Ser Val  
100 105 110  
Asp Gln Phe Val Ser Ala Ile Lys Gln Val Ala Leu Ala Asn Lys Lys  
115 120 125  
10 Trp Ile Pro Pro Gly Lys Gly Thr Leu Tyr Ile Arg Pro Ile Leu  
130 135 140  
Phe Gly Ser Gly Pro Ile Leu Gly Ser Phe Pro Ile Pro Glu Thr Thr  
145 150 155 160  
15 Phe Thr Ala Phe Ala Cys Pro Val Gly Arg Tyr His Lys Asp Asn Ser  
165 170 175  
Gly Leu Asn Leu Lys Ile Glu Asp Gln Phe Arg Arg Ala Phe Pro Ser  
180 185 190  
Gly Thr Gly Gly Val Lys Ser Ile Thr Asn Tyr Cys Pro Val Trp Ile  
195 200 205  
20 Pro Leu Ala Glu Ala Lys Lys Gln Gly Phe Ser Asp Ile Leu Phe Leu  
210 215 220  
Asp Ala Ala Thr Gly Lys Asn Ile Glu Glu Leu Phe Ala Ala Asn Val  
225 230 235  
25 Phe Met Leu Lys Gly Asn Val Val Ser Thr Thr Thr Ile Ala Gly Thr  
240 245 250 255  
Ile Leu Pro Gly Val Thr Arg Asn Cys Val Met Glu Leu Cys Arg Asp  
260 265 270  
Phe Gly Tyr Gln Val Glu Glu Arg Thr Ile Pro Leu Val Asp Phe Leu  
275 280 285  
30 Asp Ala Asp Glu Ala Phe Cys Thr Gly Thr Ala Ser Ile Val Thr Ser  
290 295 300 305  
Ile Ala Ser Val Thr Phe Lys Asp Lys Thr Gly Phe Lys Thr Gly  
310 315 320  
35 Lys Lys His Trp Leu Arg Ser Tyr Thr Arg Arg  
325 330 335  
(2) INFORMATION FOR SEQ ID NO:74:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 297 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) FEATURE:  
(ii) MOLECULE TYPE: peptide  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..297  
(D) OTHER INFORMATION: / Ceres Seq. ID 1014998  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:  
Met Tyr Val Ala Lys Cys Asn His Gly Cys Ser Phe Gln Glu Gly Lys  
1 5 10 15  
Ile Leu Pro Phe Ala Asp Leu Leu Asn Pro Cys Ala Ala Val Leu  
20 25 30  
Gln Tyr Gly Gln Gly Leu Tyr Glu Gly Leu Lys Ala Tyr Arg Thr Glu  
35 40 45  
55 Asp Gly Arg Ile Leu Leu Phe Arg Pro Asp Gln Asn Gly Leu Arg Leu  
50 55 60  
Gln Ala Gly Ala Asp Arg Leu Tyr Met Pro Tyr Pro Ser Val Asp Gln  
65 70 75 80  
Phe Val Ser Ala Ile Lys Gln Val Ala Leu Ala Asn Lys Lys Trp Ile  
85 90 95  
60 Pro Pro Pro Gly Lys Gly Thr Leu Tyr Ile Arg Pro Ile Leu Phe Gly  
100 105 110

361  
 Ser Gly Pro Ile Leu Gly Ser Phe Pro Ile Pro Glu Thr Thr Phe Thr  
 115 120 125  
 Ala Phe Ala Cys Pro Val Gly Arg Tyr His Lys Asp Asn Ser Gly Leu  
 130 135 140  
 Asn Leu Lys Ile Glu Asp Glu Phe Arg Arg Ala Phe Pro Ser Gly Thr  
 145 150 155 160  
 Gly Gly Val Lys Ser Ile Thr Asn Tyr Cys Pro Val Trp Ile Pro Leu  
 165 170 175  
 Ala Glu Ala Lys Lys Cln Gly Phe Ser Asp Ile Leu Phe Leu Asp Ala  
 180 185 190  
 Ala Thr Gly Lys Asn Ile Glu Glu Leu Phe Ala Ala Asn Val Phe Met  
 195 200 205  
 Leu Lys Gly Asn Val Val Ser Thr Pro Thr Ile Ala Gly Thr Ile Leu  
 210 215 220  
 Pro Gly Val Thr Arg Asn Cys Val Met Glu Leu Cys Arg Asp Phe Gly  
 225 230 235 240  
 Tyr Glu Val Glu Glu Arg Thr Ile Pro Leu Val Asp Phe Leu Asp Ala  
 245 250 255  
 Asp Glu Ala Phe Cys Thr Gly Thr Ala Ser Ile Val Thr Ser Ile Ala  
 260 265 270  
 Ser Val Thr Phe Lys Asp Lys Lys Thr Gly Phe Thr Gly Lys Lys  
 275 280 285  
 His Trp Leu Arg Ser Tyr Thr Arg Arg  
 290 295

25 (2) INFORMATION FOR SEQ ID NO:75:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 785 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (11) MOLECULE TYPE: DNA (genomic)  
 (1x) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..789

(D) OTHER INFORMATION: / Ceres Seq. ID 1015323

(1x) SEQUENCE DESCRIPTION: SEQ ID NO:75:

ggtagtggt tgcctcaaa acgcgcaga gaagaagaag aagtcacaga gcatacaat  
 60  
 ggcgcataa tgcataccc gattgcgat gcgcgcgta ttctctctt caatcacac  
 120  
 catcacacc gcacgcctt gcaaaacct cctaattcc tcatattcc tctcaatac  
 180  
 ccacagaacc ctaaacctc ataccgttt cactccaca cgttcgtca ccgtttcc  
 240  
 atccgtctt ctatacttc accacgtgtt tctcttctt gtaacgtc gccacgaga  
 300  
 gccacagtc caatgcgtc gtctctccc ttgatctct gataacata attcgtttg  
 360  
 agatcgacc agagatctc ttacgcgtc cgtcgcgtt ttgttcgtg ttgttcgtg  
 420  
 tctttaac gcgcataa tctacttg ttgcgtctt ttgcgtctt gtcacagta  
 480  
 tgatttcga gaagagaag atgattatga gaatgatga ttgatgctg ctacgttga  
 540  
 gaatttggt tatgttaaga ttccagctc agctcgtct cctggaag agctcgtt  
 600  
 attgaggtt gtaatttca gttctctc attatcatga ttacgttgt tgcattaca  
 660  
 ttgttcgga tacaagaaa ttgcgtctg ttatggatt atctaactt actatactt  
 720  
 agtatcagt gttttacga ttgttagctg acgtttctta tatarata taagtatct  
 780  
 tatgatgc

(2) INFORMATION FOR SEQ ID NO:76:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 180 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (11) MOLECULE TYPE: peptide  
 (1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..180

(D) OTHER INFORMATION: / Ceres Seq. ID 1015324

(1x) SEQUENCE DESCRIPTION: SEQ ID NO:76:

362  
 Met Thr Ser Ser Tyr Leu Arg Phe Ala Ile Ala Val Val Ala Phe  
 1 5 10 15  
 Leu Ser Ile Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr  
 20 25 30  
 Ile Ser Ser Tyr Ser Leu Ser Ile Thr Thr Thr Thr Thr Thr Thr  
 35 40 45  
 Ser Asp Phe Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr  
 50 55 60  
 Leu Asn Pro His His Val Val Pro Phe Phe Val Asn Arg Arg His Glu  
 65 70 75 80  
 Lys Pro Glu Ile Glu Ser Asp Arg Ser Leu Pro Leu Ile Ser Asp Asn  
 85 90 95  
 Ile Asn Ser Phe Arg Asp Arg Thr Arg Asp Ile Leu Ser Val Val Val  
 100 105 110  
 Ala Leu Leu Phe Gly Val Gly Cys Gly Ala Leu Thr Ala Thr Met  
 115 120 125  
 Tyr Leu Val Trp Ala Leu Val Val Asn Arg Gln Ser Tyr Asp Phe Glu  
 130 135 140  
 Glu Glu Glu Asp Asp Tyr Glu Asn Asp Glu Ser Asp Ala Ala Ser Leu  
 145 150 155 160  
 Lys Lys Leu Gly Tyr Val Lys Ile Pro Ala Pro Ala Pro Ala Pro Val  
 165 170 175  
 Lys Glu Ala Ala  
 180

25 (2) INFORMATION FOR SEQ ID NO:77:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 495 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (11) MOLECULE TYPE: DNA (genomic)  
 (1x) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..495

(D) OTHER INFORMATION: / Ceres Seq. ID 1016486

(1x) SEQUENCE DESCRIPTION: SEQ ID NO:77:

actctccat tagcatctca cagctgcac actatcaat tcttcgaaa tctctctca  
 60  
 tccctctaa aaatgaagt cctgtgcgt ttatctccg ctgtctctt ctgtctaat  
 120  
 gtattcatt caacaggat ggtccagtc accgtgagg cagcacatg lgaagcaag  
 180  
 aqccataagt taagggttc atgtgagc acacacact gtgcaacgt ggcacacac  
 240  
 gaaggtctg gcgaagtaa atgcgtgga ttcgcgttc gtgcacgt cacaagcac  
 300  
 tctgatcca tcatctcca tgaactaat ctgcagcca tctcagtg gttactctt  
 360  
 tctctcaa atctctcca cgtaccata tgcacgta catgagtg tctcttaata  
 420  
 agcttggtt t-gt-gtctt ccgttttaa tgaatgta atcaattaa tggctttta  
 480  
 tatattgat tatgg

(2) INFORMATION FOR SEQ ID NO:78:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 101 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (11) MOLECULE TYPE: peptide  
 (1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..101

(D) OTHER INFORMATION: / Ceres Seq. ID 1016487

(1x) SEQUENCE DESCRIPTION: SEQ ID NO:78:

Thr Ser Ser Leu Ala Phe Tyr Ser Leu His Thr Tyr His Phe Phe Arg  
 1 5 10 15  
 Asn Leu Ser Leu Ser Leu Ser Lys Met Lys Leu Ser Val Arg Phe Ile  
 20 25 30  
 Ser Ala Ala Leu Leu Leu Phe Met Val Phe Ile Ala Thr Gly Met Gly



363

35 Pro Val Thr Val Glu Ala Arg Thr Cys Glu Ser Lys Ser His Arg Phe  
50 55 60  
Lys Gly Pro Cys Val Ser Thr His Asn Cys Ala Asn Val Cys His Asn  
65 70 75 80  
Glu Gly Phe Gly Gly Lys Cys Arg Gly Phe Arg Arg Cys Tyr  
85 90 95  
Cys Thr Arg His Cys

10 (2) INFORMATION FOR SEQ ID NO:79:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 77 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide  
(B) LOCATION: 1..77

20 (D) OTHER INFORMATION: / Ceres Seq. ID 1016488

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

Met Lys Leu Ser Val Arg Phe Ile Ser Ala Ala Leu Leu Phe Met  
1 5 10 15  
Val Phe Ile Ala Thr Gly Met Gly Pro Val Thr Val Glu Ala Arg Thr  
20 25 30  
Cys Glu Ser Lys Ser His Arg Phe Lys Gly Pro Cys Val Ser Thr His  
35 40 45  
Asn Cys Ala Asn Val Cys His Asn Glu Gly Phe Gly Gly Lys Cys  
50 55 60  
Arg Gly Phe Arg Arg Arg Cys Tyr Cys Thr Arg His Cys  
65 70 75

(2) INFORMATION FOR SEQ ID NO:80:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 62 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide  
(B) LOCATION: 1..62

(D) OTHER INFORMATION: / Ceres Seq. ID 1016489

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

Met Val Phe Ile Ala Thr Gly Met Gly Pro Val Thr Val Glu Ala Arg  
1 5 10 15  
Thr Cys Glu Ser Lys Ser His Arg Phe Lys Gly Pro Cys Val Ser Thr  
20 25 30  
His Asn Cys Ala Asn Val Cys His Asn Glu Gly Phe Gly Gly Lys  
35 40 45  
Cys Arg Gly Phe Arg Arg Arg Cys Tyr Cys Thr Arg His Cys  
50 55 60

(2) INFORMATION FOR SEQ ID NO:81:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 325 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

(A) NAME/KEY: -  
(B) LOCATION: 1..325

(D) OTHER INFORMATION: / Ceres Seq. ID 1018341

364

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:  
atcatcaaac acaaacact caatcaaaa tataaatc aagtgtaag caaacaga  
60  
acgagagaa atggaacaa agcaaacgc gagctacca ggcgtcaag ccaatggca  
120  
gactaagag aagccggcg gaatgagga caagccaag gatgtgtg cttagctca  
180  
agactcctg caacaactg gacaacaat caagagaag gcscagag csgtctgt  
240  
cgtcaaggac aagcaggca tgacaagaag ccacaaaga agcatctgg actcttcc  
300  
taattaatt cctcttcca ctgt

(2) INFORMATION FOR SEQ ID NO:82:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 68 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide  
(B) LOCATION: 1..68

(D) OTHER INFORMATION: / Ceres Seq. ID 1018342

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

Met Asp Asn Lys Gln Asn Ala Ser Tyr Gln Ala Gly Gln Ala Thr Gly  
1 5 10 15  
Gln Thr Lys Glu Lys Ala Gly Gly Met Met Asp Lys Ala Lys Asp Ala  
20 25 30  
Ala Ala Ser Ala Gln Asp Ser Leu Gln Gln Thr Gly Gln Gln Met Lys  
35 40 45  
Glu Lys Ala Gln Gly Ala Ala Asp Val Val Lys Asp Lys Thr Gly Met  
50 55 60  
Asn Lys Ser His

(2) INFORMATION FOR SEQ ID NO:83:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 44 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide  
(B) LOCATION: 1..44

(D) OTHER INFORMATION: / Ceres Seq. ID 1018343

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

Met Met Asp Lys Ala Lys Asp Ala Ala Ser Ala Gln Asp Ser Leu  
1 5 10 15  
Gln Gln Thr Gly Gln Gln Met Lys Glu Lys Ala Gln Gly Ala Ala Asp  
20 25 30  
Val Val Lys Asp Lys Thr Gly Met Asn Lys Ser His

(2) INFORMATION FOR SEQ ID NO:84:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 43 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide  
(B) LOCATION: 1..43

(D) OTHER INFORMATION: / Ceres Seq. ID 1018344

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

Met Asp Lys Ala Lys Asp Ala Ala Ser Ala Gln Asp Ser Leu Gln  
1 5 10 15  
Gln Thr Gly Gln Gln Met Lys Glu Lys Ala Gln Gly Ala Ala Asp Val

365

Val Lys Asp Lys Thr Gly Met Asn Lys Ser His  
35 40 30

(2) INFORMATION FOR SEQ ID NO:85:

- (A) LENGTH: 540 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(E) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..540  
(D) OTHER INFORMATION: / Ceres Seq. ID 1018382

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:  
taaaagcaat aaaaacaaa aatggcggtt actccgaaga tcaatcaatg cctcattgtc 60  
cttaagatct acatgacatc cccaacagay tcaucacatc agtctgggac agtgcagagc 120  
acactggcac agtctgggac ctacttgacc aacagtggtc cactggcacc acaatgtcgc 180  
gtggagtcac agtcttggtg ccaattggct cagaccacac cggagcgtaa acaatgtcgc 240  
gagtgctta aac-agcggg taaagaatcc aagggcctca acaccgacct tggcgcgca 300  
ctctcaca cttgggtgtg tcaatccc taccacatca gttttagcac caatggagac 360  
agatataga ctggcggtg aagaaggcta gtatcagat gtaagactaa tcaaatgtgc 420  
cagcttttaa cctaatkaa taaagatc ctgtttatat tcccatatt atgttttat 480  
ctcttatct atgaacacac acgatttcat atgtaataa tgacaacgga tctttctctc 540

(2) INFORMATION FOR SEQ ID NO:86:

- (A) LENGTH: 119 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(E) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..119  
(D) OTHER INFORMATION: / Ceres Seq. ID 1018383

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

Met Ala Phe Thr Pro Lys Ile Ile Thr Cys Leu Ile Val Leu Thr Ile  
1 5 10 15  
Tyr Met Thr Ser Pro Thr Glu Ser Thr Ile Gln Cys Gly Thr Val Thr  
20 25 30  
Ser Thr Leu Ala Gln Cys Val Thr Tyr Ileu Thr Asn Ser Gly Pro Leu  
35 40 45  
Pro Ser Gln Cys Cys Val Gly Val Lys Ser Leu Tyr Gln Leu Ala Gln  
50 55 60  
Thr Thr Pro Asp Arg Lys Gln Val Cys Glu Cys Leu Lys Leu Ala Gly  
65 70 75 80  
Lys Glu Ile Lys Gly Leu Asn Thr Asp Leu Val Ala Ala Leu Pro Thr  
85 90 95  
Thr Cys Gly Val Ser Ile Pro Tyr Pro Ile Ser Phe Ser Thr Asn Cys  
100 105 110  
Asp Ser Ile Ser Thr Ala Val  
115

(2) INFORMATION FOR SEQ ID NO:87:

- (A) LENGTH: 102 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(E) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..102

366

(D) OTHER INFORMATION: / Ceres Seq. ID 1018384

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

Met Thr Ser Pro Thr Glu Ser Thr Ile Gln Cys Gly Thr Val Thr Ser  
1 5 10 15  
Thr Leu Ala Gln Cys Val Thr Tyr Leu Thr Asn Ser Gly Pro Leu Pro  
20 25 30  
Ser Gln Cys Cys Val Gly Val Lys Ser Leu Tyr Gln Leu Ala Gln Thr  
35 40 45  
Thr Pro Asp Arg Lys Gln Val Cys Glu Cys Leu Lys Leu Ala Gly Lys  
50 55 60  
Glu Ile Lys Gly Leu Asn Thr Asp Leu Val Ala Ala Leu Pro Thr Thr  
65 70 75 80  
Cys Gly Val Ser Ile Pro Tyr Pro Ile Ser Phe Ser Thr Asn Cys Asp  
85 90 95  
Ser Ile Ser Thr Ala Val  
100

(2) INFORMATION FOR SEQ ID NO:88:

- (A) LENGTH: 75 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(E) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..75  
(D) OTHER INFORMATION: / Ceres Seq. ID 1018385

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

Lys Ala Ile Lys Asn Lys Asn Gly Val Tyr Ser Glu Asp His His Met  
1 5 10 15  
Pro His Cys Pro Tyr Asp Leu His Asp Ile Pro Asn Arg Val Asn His  
20 25 30  
Pro Val Thr Asp Ser Asp Glu His Thr Gly Thr Val Arg Asp Leu Leu  
35 40 45  
Asp Gln Gln Thr Ser Ile Ala Ile Thr Met Leu Arg Gly Ser Gln Val  
50 55 60  
Ile Val Pro Ile Gly Ser Asp His Thr Gly Pro  
65 70 75

(2) INFORMATION FOR SEQ ID NO:89:

- (A) LENGTH: 844 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(E) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..844  
(D) OTHER INFORMATION: / Ceres Seq. ID 1020666

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

ctaaatatt aaagctttcc ttcaacgca ttgtattgcc tccctggaac accgaagctc 60  
tggataacc tattttccc ttctaaagt ctctgtgttt ctattgtatc tgtgaagatg 120  
tattcggctc aqaacaagat ccaacagat aaggtgtttg caccacaga gtctgaagag 180  
caagttactc aggtcttgtt tgacttggag aacacacac aggaattgaa aggaagattg 240  
aaagattctt acattaacca agctgttcag atgtatttt ctggcaaccg caagctgtt 300  
gtgatctag ttccattcag cgtgagaaa gctttccgca agattctatc tctctgttc 360  
agagagcttg aqaagaagt cagtgaaaa gatgtgatc ttgtctac cagaagaatc 420  
atcgccacac caagaagaag ctacgtttt cagaacacac gcaacagac tctactctt 480  
gtccatgaag ccatgtccca ggaatgcct taccctgtg agattgtgg aaagaagc 540  
agataccatc ttgatgtcac caagatattt aagttattt tggatagtaa gtaagaagc 600  
gacactgagt acacgtcaga gacactggt gggtgtaca gaaacttac aggaagaat 660  
gtagtctcag agtaccagct catagaagct tgaagaagat aatgtttgtt tcatctttt 720

367

ttgtcgata gagagctttt gattctgtt ggagtaaaa cttcttgcaa agatttgtt  
atttcagcat ttgcgaatct cttctcaaa gtgcgattta tgcatacaat tctttttcc  
tacc

(2) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 191 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..191

(D) OTHER INFORMATION: / Ceres Seq. ID 1020667

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

Met Tyr Ser Gly Gln Asn Lys Ile His Lys Asp Lys Gly Val Ala Pro  
1 10 15

Thr Glu Phe Glu Gln Gln Val Thr Gln Ala Leu Phe Asp Leu Glu Asn  
20 25 30

Thr Asn Gln Glu Leu Lys Ser Glu Leu Lys Asp Leu Tyr Ile Asn Gln  
35 40 45

Ala Val Gln Met Asp Ile Ser Gly Asn Arg Lys Ala Val Val Ile Tyr  
50 55 60

Val Pro Phe Arg Leu Arg Lys Ala Phe Arg Lys Ile His Leu Arg Leu  
65 70 75 80

Val Arg Glu Leu Glu Lys Lys Phe Ser Gly Lys Asp Val Ile Phe Val  
85 90 95

Ala Thr Arg Arg Ile Met Arg Pro Lys Lys Gly Ser Ala Val Gln  
100 105 110

Arg Pro Arg Asn Arg Thr Leu Thr Ser Val His Glu Ala Met Leu Glu  
115 120 125

Asp Val Ala Tyr Pro Ala Glu Ile Val Gly Lys Arg Thr Arg Tyr Arg  
130 135 140

Leu Asp Gly Thr Lys Ile Met Lys Val Phe Leu Asp Ser Lys Leu Lys  
145 150 155 160

Asn Asp Thr Glu Tyr Lys Leu Glu Thr Met Val Gly Val Tyr Arg Lys  
165 170 175

Leu Thr Gly Lys Asp Val Val Phe Glu Tyr Pro Val Ile Glu Ala  
180 185 190

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 140 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..140

(D) OTHER INFORMATION: / Ceres Seq. ID 1020668

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

Met Asp Ile Ser Gly Asn Arg Lys Ala Val Val Ile Tyr Val Pro Phe  
1 5 10 15

Arg Leu Arg Lys Ala Phe Arg Lys Ile His Leu Arg Leu Val Arg Glu  
20 25 30

Leu Glu Lys Lys Phe Ser Gly Lys Asp Val Ile Phe Val Ala Thr Arg  
35 40 45

Arg Ile Met Arg Pro Lys Lys Gly Ser Ala Val Gln Arg Pro Arg  
50 55 60

Asn Arg Thr Leu Thr Ser Val His Glu Ala Met Leu Glu Asp Val Ala  
65 70 75 80

Tyr Pro Ala Glu Ile Val Gly Lys Arg Thr Arg Tyr Arg Leu Asp Gly  
85 90 95

368

Thr Lys Ile Met Lys Val Phe Leu Asp Ser Lys Leu Lys Asn Asp Thr  
100 105 110 115

Glu Tyr Lys Leu Glu Thr Met Val Gly Val Tyr Arg Lys Leu Thr Gly  
120 125 130

Lys Asp Val Val Phe Glu Tyr Pro Val Ile Glu Ala  
135 140 145

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 789 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..789

(D) OTHER INFORMATION: / Ceres Seq. ID 1020784

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

aggtttaggt ttgtctcaa aatctcaga agagaaga gaagtaaa acgatacaa  
100 105 110 115 120 125 130 135 140 145

tgatctatc ategtaacct cgtattgga tgcgcgtgt agctttctc tcaatacaa  
150 155 160 165 170 175 180 185 190 195

ccatcacac cgcgcgtct tgcataact tctaatctc ctcatattc ctcatacaa  
200 205 210 215 220 225 230 235 240 245

cccaagaaa ccttaacct caatcgatt tcaactcac acgattcac accgttcaa  
250 255 260 265 270 275 280 285 290 295

caatcgctg tctaaacct caccacgtg tctctttct cgtcaactg cgtcaagaa  
300 305 310 315 320 325 330 335 340 345

agccacgat ccaatcgat cgtctcttc ctttgatct tgaacatc aattcgtta  
350 355 360 365 370 375 380 385 390 395

gagatcaac cagagatatt atagcgtg tgcgcgtgt ttgttgtgt gtttgtgtg  
400 405 410 415 420 425 430 435 440 445

gtgtttaac cgtcttact atgtattgt ttggcgtct tgttgtgt cgtcaagct  
450 455 460 465 470 475 480 485 490 495

agatttga ggaagaaga gatgattatg agaatatga atctatgct gctagtga  
500 505 510 515 520 525 530 535 540 545

gaattgagg tgaatttc agttctga tctatcatg atttctgtg tctgttctc  
550 555 560 565 570 575 580 585 590 595

atgtgtgac gtacaaatga attcgtctt gtttgatt tctcaact tctatactc  
600 605 610 615 620 625 630 635 640 645

taagtacag tgtttttacg atttgatt gctgtttct aratatat ataatgtga  
650 655 660 665 670 675 680 685 690 695

tctatgac  
700 705 710 715 720 725 730 735 740 745

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 180 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..180

(D) OTHER INFORMATION: / Ceres Seq. ID 1020785

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

Met Thr Ser Ser Tyr Leu Arg Phe Ala Ile Ala Val Val Ala Phe  
1 5 10 15 20 25 30 35 40 45

Leu Ser Ile Thr Thr Ile Thr Thr Ala Arg Pro Cys Lys Thr Phe Leu  
50 55 60 65 70 75 80 85 90 95

Ile Ser Ser Tyr Ser Leu Ser Ile Thr Pro Glu Asn Pro Asn Leu Glu  
100 105 110 115 120 125 130 135 140 145

Ser Asp Phe Thr Ser Thr Arg Phe Ile Thr Val Phe Thr Ile Arg Arg  
150 155 160 165 170 175 180 185 190 195

Leu Asn Pro His His Val Val Pro Phe Phe Val Asn Arg Arg His Glu  
200 205 210 215 220 225 230 235 240 245

Lys Pro Gln Ile Gln Ser Asp Arg Ser Leu Pro Leu Ile Ser Asp Asn  
250 255 260 265 270 275 280 285 290 295

Ile Asn Ser Phe Arg Asp Arg Thr Arg Asp Ile Leu Ser Val Val  
300 305 310 315 320 325 330 335 340 345

Ala Leu Leu Phe Gly Val Gly Cys Gly Ala Leu Thr Ala Thr Met  
350 355 360 365 370 375 380 385 390 395

Tyr Leu Val Trp Ala Leu Val Val Asn Arg Gln Ser Tyr Asp Phe Glu  
400 405 410 415 420 425 430 435 440 445

369 130 135 140 145 150 155 160 165 170 175 180

Glu Glu Glu Asp Asp Tyr Tyr Glu Asp Glu Ser Asp Ala Ala Ser Leu  
Lys Lys Leu Gly Tyr Val Lys Ile Pro Ala Pro Ala Pro Ala Pro Val  
Lys Glu Ala Ala

(2) INFORMATION FOR SEQ ID NO:94:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 755 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..765  
(D) OTHER INFORMATION: / Ceres Seq. ID 1021525

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:  
acacccactc acttttcagg ttccaccat tagcacaaa aatcgaagt ccgacacaaa 60  
attggaac acgatggg agaatcttc agaaacttc ctccaaatc ccgactactc 120  
ctccatcac cagactcta atcgttcga catcgcac gcctgctcct tctcgtctc 180  
tgggttcga gaaatcttc gccgcagcc tcaagtttc tgcgtgat ccctcaggt 240  
acaagcaga catattga ctctcggc agactcctc tgcgtgctc tccatcagc 300  
gctgataga tccagatgg catagatgg atgacacga ggttcgcca gcgagtgcc 360  
aggttcgat tccagagga ttgctcaga agctccgca ggacttac gatgaggt 420  
atgctcaga gcgaactc agatccgta tcttgacaa gcaactcgt cttggttgc 480  
agttgctt aacgcagg ctccaggg ctctcagg tgcctcgc tgcctcga gataagctt 540  
ggatctcc gaagtttt tctctagg atccataaa tcttggttc gttcctaat 600  
tgcataaa aatcgttt ggaatttc gtaactga gatactca ttgaattc 660  
taatttgc ttgtctac taatttgc taatttgc taatttgc taatttgc 720  
cgtctcgt cttcaagt ggtgaagt tctttctg atcc

(2) INFORMATION FOR SEQ ID NO:95:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 159 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..159  
(D) OTHER INFORMATION: / Ceres Seq. ID 1021526

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:  
Met Ala Thr Thr Leu Glu Lys Leu Ser Ser Glu Ile His Arg Leu Ser 1  
Pro Phe Thr Arg Ser Leu Ile Val Arg Thr Ser Ala Thr Ser Ala Pro 5  
Ser Pro Ser Leu Gly Ser Lys Lys Val Ser Asp Arg Ile Val Lys Leu 10  
Ser Ala Ile Asp Pro Asp Gly Tyr Lys Glu Asp Ile Ile Gly Leu Ser 15  
Gly Glu Thr Leu Leu Arg Ala Leu Thr His Thr Gly Leu Ile Asp Pro 20  
Ala Ser His Arg Leu Asp Asp Ile Glu Ala Cys Ser Ala Glu Cys Glu 25  
Val Glu Ile Ala Glu Thr Lys Leu Lys Leu Pro Pro Arg Thr Tyr 30  
Asp Glu Glu Tyr Val Leu Lys Arg Ser Ser Arg Ile Leu Asn 35  
Lys His Ser Arg Leu Gly Cys Glu Val Val Leu Thr Glu Glu Leu Glu 40

370 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 900 905 910 915 920 925 930 935 940 945 950 955 960 965 970 975 980 985 990 995

Gly Met Val Val Ala Val Pro Glu Ala Lys Pro Trp Asp Ile Pro  
(2) INFORMATION FOR SEQ ID NO:96:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 538 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..588  
(D) OTHER INFORMATION: / Ceres Seq. ID 1021563

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:  
attctggaa agtctcatt ctcgacccc aatctgaga ttagggtta agaacatt 60  
tttattcgg tgcgcaaa acaatcccg atcgaaaag gaagaaga tcgaatggc 120  
tttgaggg gttacagtg aatcaggg gaagaagt agcagcttc caggcttat 180  
caatcagc ttctaatg agacctga gacctctg agcagggc tcgatacta 240  
caacgaaa tcaatcga ctagctcgt tgcctatc cttcattc gctttcgg 300  
catgcttc tctacatt tgcctatc taatgagc tgcctatc tgcctatc 360  
gccttaag gacagctg gctatgac tgcgggaa cgtttcgt cgcctatc 420  
tttggggg ttgctgaa atcttctc gctcaggt gacagact caagattgt 480  
gcttattg tctctctt gaatttcc tgcctatc tgcctatc tgcctatc 540  
cttcaagt acatcctc atagccta atcgagct tgcctatc tgcctatc

(2) INFORMATION FOR SEQ ID NO:97:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 128 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..128  
(D) OTHER INFORMATION: / Ceres Seq. ID 1021564

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:  
Ser Ser Glu Ser Leu Ile Ser Arg Ser Pro Ile Arg Gly Leu Gly Leu 1  
Lys Glu Pro Phe Leu Phe Ser Ser Arg Asn Asn Lys Ser Arg Ser Lys 5  
Lys Glu Glu Glu Ile Glu Met Ala Leu Arg Arg Val Tyr Ser Glu Ile 10  
Arg Gly Lys Lys Val Thr Glu Leu Pro Gly Tyr Ile Lys Ser Thr Phe 15  
Ser Met Glu Thr Val Lys Thr Ser Val Lys Arg Gly Leu Asp Asn Tyr 20  
Asn Glu Lys Tyr Ile Glu Thr Ser Ser Val Asp Pro Ile Leu His Ile 25  
Cys Phe Tyr Gly Met Ala Phe Ser Tyr Leu Val Ala Leu Pro Asn Glu 30  
Arg Arg His Leu Glu His Glu Glu His Ala Lys Glu His Gly Gly His 35  
(2) INFORMATION FOR SEQ ID NO:98:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 90 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..90

371

(D) OTHER INFORMATION: / Ceres Seq. ID 1021565

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

Met Ala Leu Arg Arg Val Tyr Ser Glu Ile Arg Lys Lys Val Thr  
1 5 10 15  
Glu Leu Pro Gly Tyr Ile Lys Ser Thr Phe Ser Met Glu Thr Val Lys  
20 25 30  
Thr Ser Val Lys Arg Gly Leu Asp Asn Tyr Asn Glu Lys Tyr Ile Glu  
35 40 45  
Thr Ser Ser Val Asp Pro Ile Leu His Ile Cys Phe Tyr Gly Met Ala  
50 55 60  
Phe Ser Tyr Leu Val Ala Leu Pro Asn Glu Arg Arg His Leu Glu His  
65 70 75 80  
Gln Gln His Ala Lys Glu His Gly Gly His  
85 90

(2) INFORMATION FOR SEQ ID NO:99:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 586 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..586

(D) OTHER INFORMATION: / Ceres Seq. ID 1021576

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

ttgcctaaat cattttaaag actgtataga gaaacaaa actgcacaaa -aaaaataa  
60  
aaaaacatc cacaagaaa taagaattt gtaaatcaa ctaagaaaat ggcagcact  
120  
atgatgacta catgtctcca gttaacaggt cttagagcca caaaatttc tgcagctct  
180  
gtacaagcc tggcaagctt tgaacagag agacgaaggg gaaatggagc ttgggtgca  
240  
aagtggact tcatcggttc attcaacaa ctgataatgg taacgtgac gacctggag  
300  
ttgttcggt ggaattcgg acttgccca tcaagcaata gaaagcgac agctggactt  
360  
aggttgagc cagtgactc agctctaca acaggtgacc cagcaggtt cagcttgg  
420  
gacatttgg gtcttgacc cgttgctc: atcatcgtt taggtattgt ccttggcctt  
480  
aaaaacattg gtctatttgc agttctcaa agctcttta ttgtatttg taaaattgt  
540  
agattttat aacaatttc tcatgacc: gaacgagatc taatgc

(2) INFORMATION FOR SEQ ID NO:100:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 130 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: -

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..130

(D) OTHER INFORMATION: / Ceres Seq. ID 1021577

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

Met Ala Ser Thr Met Thr Thr Leu Pro Gln Phe Asn Gly Leu Arg  
1 5 10 15  
Ala Thr Lys Ile Ser Ala Ala Pro Val Gln Gly Leu Ala Ser Val Gln  
20 25 30  
Pro Met Arg Arg Lys Gly Asn Gly Ala Leu Gly Ala Lys Cys Asp Phe  
35 40 45  
Ile Gly Ser Ser Thr Asn Leu Ile Met: Val Thr Ser Thr Thr Leu Met  
50 55  
Leu Phe Ala Gly Arg Phe Gly Leu Ala Pro Ser Ala Asn Arg Lys Ala  
60 65 70 75 80  
Thr Ala Gly Leu Arg Leu Glu Ala Arg Asp Ser Gly Leu Gln Thr Gly  
85 90 95  
Asp Pro Ala Gly Phe Thr Leu Ala Asp Thr Leu Ala Cys Gly Thr Val  
100 105 110

372

Gly His Ile Ile Gly Val Gly Val Val Leu Gly Leu Lys Asn Ile Gly  
115 120 125  
Ala Ile  
130

(2) INFORMATION FOR SEQ ID NO:101:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 126 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..126

(D) OTHER INFORMATION: / Ceres Seq. ID 1021578

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

Met Met Thr Thr Leu Pro Gln Phe Asn Gly Leu Arg Ala Thr Lys Ile  
1 5 10 15  
Ser Ala Ala Pro Val Gln Gly Leu Ala Ser Val Gln Pro Met Arg Arg  
20 25 30  
Lys Gly Asn Gly Ala Leu Gly Ala Lys Cys Asp Phe Ile Gly Ser Ser  
35 40 45  
Thr Asn Leu Ile Met Val Thr Ser Thr Thr Leu Met Leu Phe Ala Gly  
50 55 60  
Arg Phe Gly Leu Ala Pro Ser Ala Asn Arg Lys Ala Thr Ala Gly Leu  
65 70 75 80  
Arg Leu Glu Ala Arg Asp Ser Gly Leu Gln Thr Gly Asp Pro Ala Gly  
85 90 95  
Phe Thr Leu Ala Asp Thr Leu Ala Cys Gly Thr Val Gly His Ile Ile  
100 105 110  
Gly Val Gly Val Val Leu Gly Leu Lys Asn Ile Gly Ala Ile  
115 120 125

(2) INFORMATION FOR SEQ ID NO:102:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 125 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..125

(D) OTHER INFORMATION: / Ceres Seq. ID 1021579

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

Met Thr Thr Leu Pro Gln Phe Asn Gly Leu Arg Ala Thr Lys Ile Ser  
1 5 10 15  
Ala Ala Pro Val Gln Gly Leu Ala Ser Val Gln Pro Met Arg Arg Lys  
20 25 30  
Gly Asn Gly Ala Leu Gly Ala Lys Cys Asp Phe Ile Gly Ser Ser Thr  
35 40 45  
Asn Leu Ile Met Val Thr Ser Thr Thr Leu Met Leu Phe Ala Gly Arg  
50 55 60  
Phe Gly Leu Ala Pro Ser Ala Asn Arg Lys Ala Thr Ala Gly Leu Arg  
65 70 75 80  
Leu Glu Ala Arg Asp Ser Gly Leu Gln Thr Gly Asp Pro Ala Gly Phe  
85 90 95  
Thr Leu Ala Asp Thr Leu Ala Cys Gly Thr Val Gly His Ile Ile Gly  
100 105 110  
Val Gly Val Val Leu Gly Leu Lys Asn Ile Gly Ala Ile  
115 120 125

(2) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:

373

- (A) LENGTH: 801 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ix) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..801

(D) OTHER INFORMATION: / Ceres Seq. ID 1021927

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

attttctctt cagatattccat aaaaagagag agagaataat aaagagaana ctgaagaagc 60  
 tagagatgg agaaagatga tgaatcagac agactagacc cggtaggtgt 120  
 gccacggaga agtgagaga gactgagctc ggaatccgaa ctgcagagac aatgctcgg 180  
 ttgactcggg tgggtctttg tttgacacg ctgtttctca ttttaaga cttcgaact 240  
 aaagattcgg gttaatttc tttctccat ctacacagct ttggacctt ggtcagcca 300  
 aa-ggaatat gtgcagatca cttctctca tccagaccca ttgcagcat gctcgtttt 360  
 tcttgacaa tgcctcgtgt ttggaccttc ttgtgctcg accagcttct gactacatg 420  
 qttctgctg ctgagactgt acacagcgag gttctatact tggctacaa tggagactca 480  
 qcattactt ggcagatag actgactccc taaggcgggt ttgtcatag agccactgt 540  
 tctgtataa tcaattctt ttggtttgt ttctacatgt ttctctctt aatctctct 600  
 taagactct taactcgtt tgatctctt tccattgtg acttcgcaa gaattctgaa 660  
 gtcgtgtct tggagatga gactctcat ttgctctca agtacctata gttgtcctg 720  
 tctcaaatg ttgtctttt taacttgggt caagagaag aatgcttatg tattctctct 780  
 tgttcattg attttctctt c

(2) INFORMATION FOR SEQ ID NO:104:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 204 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ix) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..204

(D) OTHER INFORMATION: / Ceres Seq. ID 1021928

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

Met Glu Lys Ser Asn Asp His Asp Lys Ala Ser His Gly Gly Ser Gly 15  
 1 Gly Gly Ala Thr Glu Lys Trp Glu Glu Thr Ser Leu Gly Ile Arg Thr 20  
 Ala Glu Thr Met Leu Asg Leu Ala Pro Val Gly Leu Cys Val Ala Ala 25  
 35 Leu Val Val Met Leu Lys Asp Ser Glu Thr Asn Glu Phe Gly Ser Ile 50  
 Ser Tyr Ser Asn Leu Thr Ala Phe Arg Tyr Leu Val His Ala Asn Gly 55  
 65 Ile Cys Ala Gly Tyr Ser Leu Leu Ser Ala Ala Ile Ala Ala Met Pro 70  
 Arg Ser Ser Ser Met Pro Arg Val Trp Thr Phe Phe Cys Leu Asp 75  
 100 105 110 115  
 50 Glu Leu Leu Thr Tyr Leu Val Leu Ala Ala Gly Ala Val Ser Ala Glu 120  
 Val Leu Tyr Leu Ala Tyr Asn Gly Asp Ser Ala Ile Thr Trp Ser Asp 125  
 130 135 140 145  
 55 Ala Cys Ser Ser Tyr Gly Gly Phe Cys His Arg Ala Thr Ala Ser Val 150  
 Ile Ile Thr Phe Phe Val Val Cys Phe Tyr Ile Val Leu Ser Leu Ile 155  
 160 165 170 175  
 60 Ser Ser Tyr Lys Leu Phe Thr Arg Phe Asp Pro Ser Ile Val Asp 180  
 Ser Ala Lys Asn Leu Glu Val Ala Val Phe Gly Ser 185  
 190 195 200

374

- (2) INFORMATION FOR SEQ ID NO:105:  
 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 169 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ix) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..169

(D) OTHER INFORMATION: / Ceres Seq. ID 1021929

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

Met Leu Arg Leu Ala Pro Val Gly Leu Cys Val Ala Ala Leu Val Val 15  
 1 Met Leu Lys Asp Ser Glu Thr Asn Glu Phe Gly Ser Ile Ser Tyr Ser 20  
 Asn Leu Thr Ala Phe Arg Tyr Leu Val His Ala Asn Gly Ile Cys Ala 25  
 30 Gly Tyr Ser Leu Leu Ser Ala Ala Ile Ala Ala Met Pro Arg Ser Ser 35  
 40 Ser Thr Met Pro Arg Val Trp Thr Phe Phe Cys Leu Asp Gln Leu Leu 45  
 50 55 60 65 70 75 80  
 Thr Tyr Leu Val Leu Ala Ala Gly Ala Val Ser Ala Glu Val Leu Tyr 85  
 90 95  
 25 Leu Ala Tyr Asn Gly Asp Ser Ala Ile Thr Trp Ser Asp Ala Cys Ser 100  
 Ser Tyr Gly Gly Phe Cys His Arg Ala Thr Ala Ser Val Ile Ile Thr 105  
 110 115 120 125  
 30 Phe Phe Val Val Cys Phe Tyr Ile Val Leu Ser Leu Ile Ser Ser Tyr 130  
 135 140  
 Lys Leu Phe Thr Arg Phe Asp Pro Ser Ile Val Asp Ser Ala Lys 145  
 150 155  
 Asn Leu Glu Val Ala Val Phe Gly Ser 160  
 165

(2) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 153 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ix) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..153

(D) OTHER INFORMATION: / Ceres Seq. ID 1021930

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

Met Leu Lys Asp Ser Glu Thr Asn Glu Phe Gly Ser Ile Ser Tyr Ser 15  
 1 Asn Leu Thr Ala Phe Arg Tyr Leu Val His Ala Asn Gly Ile Cys Ala 20  
 25 Gly Tyr Ser Leu Leu Ser Ala Ala Ile Ala Ala Met Pro Arg Ser Ser 30  
 35 40 45  
 Ser Thr Met Pro Arg Val Trp Thr Phe Phe Cys Leu Asp Gln Leu Leu 50  
 55 60 65  
 Thr Tyr Leu Val Leu Ala Ala Gly Ala Val Ser Ala Glu Val Leu Tyr 70  
 75 80  
 Leu Ala Tyr Asn Gly Asp Ser Ala Ile Thr Trp Ser Asp Ala Cys Ser 85  
 90 95  
 Ser Tyr Gly Gly Phe Cys His Arg Ala Thr Ala Ser Val Ile Ile Thr 100  
 105 110  
 Phe Phe Val Val Cys Phe Tyr Ile Val Leu Ser Leu Ile Ser Ser Tyr 115  
 120 125

375

Lys Leu Phe Thr Arg Phe Asp Pro Pro Ser Ile Val Asp Ser Ala Lys  
130 135 140  
Asn Leu Glu Val Ala Val Phe Gly Ser  
145 150

5 (2) INFORMATION FOR SEQ ID NO:107:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 602 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..602

(D) OTHER INFORMATION: / Ceres Seq. ID 1021945

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

atcagcgat ggattggca gaactggg cgaattcgg acccgattc tccggccg  
tcttgaac cgggtggtg gtcagacc gtcattgca gttcatcca agtccctc  
gtccatcc tccgggat attcctct ctcggagctt tgatttcaa tgggtcaga  
aagaadaca ttgattacc tcttaccg gaggcgagt ggagattgaa gctgggctt  
ttcatagct atgcgtac attgttcc ttactgctt ctatggcgt gctgattcaa  
gattcgttg tgaagactgg gcttcaact tggactggg tggctggct cttcaagt  
gaatttat tgaagactgg gctaatgat tggacatcg accagagta ggcacaggc  
acaaatcca tctccagctc agattcatc ttgacaact tacaagttc tctgttct  
ctaactgta attcctgtg tctgttgg tcaactggc algtgtgcg agattcctt  
gtgcatcag tttgtttc tgttaaat tttgtgagt gttattgaa atattccaa  
cc

20 (2) INFORMATION FOR SEQ ID NO:108:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 136 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..136

(D) OTHER INFORMATION: / Ceres Seq. ID 1021946

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

Ile Glu Arg Trp Ile Trp Gln Asn Cys Gly Arg Phe Ser Asp Pro Asp  
1 5 10 15  
Ser Pro Ala Pro Ser Ser Glu Pro Gly Gly Val Asp Ala Val  
20 25 30  
Cys Ser Ser Ile Gln Val Pro Phe Val His Tyr Leu Pro Gly Ile Phe  
35 40 45  
Ala Ser Leu Gly Ala Leu Met Phe Asn Cys Val Arg Lys Glu Asp Ile  
50 55 60  
Asp Tyr Ser Pro Tyr Asp Glu Gly Glu Trp Arg Leu Lys Leu Trp Leu  
65 70 75 80  
Phe Ile Ala Tyr Val Val Ala Phe Val Ser Leu Ala Ala Ser Val Gly  
85 90 95  
Leu Leu Ile Gln Asp Ser Val Val Lys Thr Gly Pro Ser Thr Trp Thr  
100 105 110  
Gly Val Ala Gly Val Phe Gln Cys Val Phe Val Leu Ile Ser Gly Leu  
115 120 125  
Met Tyr Trp Thr Ser His Ser Glu  
130 135

55 (2) INFORMATION FOR SEQ ID NO:109:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 82 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

376

(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..82

(D) OTHER INFORMATION: / Ceres Seq. ID 1021947

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

Met Phe Asn Cys Val Arg Lys Glu Asp Ile Asp Tyr Ser Pro Tyr Asp  
1 5 10 15  
Glu Gly Glu Trp Arg Leu Lys Leu Trp Phe Ile Ala Tyr Val Val  
20 25 30  
Ala Phe Val Ser Leu Ala Ala Ser Val Gly Leu Ile Gln Asp Ser  
35 40 45  
Val Val Lys Thr Gly Pro Ser Thr Trp Thr Gly Val Ala Gly Val Phe  
50 55 60  
Gln Cys Val Phe Val Leu Ile Ser Gly Leu Met Tyr Trp Thr Ser His  
65 70 75 80  
Ser Glu

20 (2) INFORMATION FOR SEQ ID NO:110:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 693 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..693

(D) OTHER INFORMATION: / Ceres Seq. ID 1022170

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

gagtcgctt tcttcttc cscatttt ttcttctag gtcgaggt taaagagaa  
ggtttcaat taggtttg tagagaaa gttggccga ag tgggaa taccgtgaa  
gtttttcac gaggtctag gtcatacgt gaggtggag cttaaaggag ccaggttta  
cagaggaagt agatttqat ttgaggttaa ttggaactgt cagctcagg atattctta  
tccgcagg gttggcagg tatcagct ttgacatgc ttacttcag cagataaggt  
caggtttat gtcataccag acatttcaa acagtcca atgttcaag gttatgagc  
tagatacag ggaagagct catcactgg tttggcaga gtagagctg caatgcgag  
gaaccggct ccgggccc ggcgtgaa tggaggaag gtagagctac cactctgag  
gagatgat atttactc caactgaa gctttgattc tggatagat tttctgat  
accaggtga aagccgctt tctgtttt gtcattaatg ctttagagaa cactgtaaa  
cacaacaga gagagatara tccagatgca ctgagtgtc tgtatttgg gttaccaaa  
-gtttgaatc attgacaagt tttcgttgc atc

45 (2) INFORMATION FOR SEQ ID NO:111:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 131 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..131

(D) OTHER INFORMATION: / Ceres Seq. ID 1022171

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

Met Ser Arg Ser Leu Gly Ile Pro Val Lys Leu Leu His Gln Ala Ser  
1 5 10 15  
Gly His Ile Val Thr Val Glu Leu Lys Ser Gly Glu Leu Tyr Arg Gly  
20 25 30  
Ser Met Ile Glu Cys Glu Asp Asn Trp Asn Cys Gln Leu Gln Asp Ile  
35 40 45  
Thr Tyr Thr Ala Lys Asp Gly Lys Val Ser Gln Leu Glu His Val Phe

377

50 Ile Arg Gly Ser Lys Val Arg Phe Met Val Ile Pro Asp Ile Leu Lys 60  
 55 His Ala Pro Met Phe Lys Arg Leu Asp Ala Arg Ile Lys Gly Lys Ser 75  
 80 Ser Ser Leu Gly Val Gly Arg Gly Arg Ala Ala Met Arg Gly Lys Pro 85  
 100 Ala Gly Pro Gly Arg Gly Thr Gly Arg Gly Ala Val Pro Pro 110  
 115 Val Arg Arg 120  
 130

(2) INFORMATION FOR SEQ ID NO:112:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 98 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..98

(D) OTHER INFORMATION: / Ceres Seq. ID 10221/2

(1) SEQUENCE DESCRIPTION: SEQ ID NO:112:

25 Met Ile Glu Cys Glu Asp Asn Trp Asn Cys Glu Leu Glu Asp Ile Thr 10  
 Tyr Thr Ala Lys Asp Gly Lys Val Ser Glu Leu Glu His Val Phe Ile 15  
 Arg Gly Ser Lys Val Arg Phe Met Val Ile Pro Asp Ile Leu Lys His 20  
 30 Ala Pro Met Phe Lys Arg Leu Asp Ala Arg Ile Lys Gly Lys Ser Ser 25  
 Ser Leu Gly Val Gly Arg Gly Arg Ala Ala Met Arg Gly Lys Pro Ala 30  
 35 Ala Gly Pro Gly Arg Gly Thr Gly Gly Arg Gly Ala Val Pro Pro Val 35  
 Arg Arg 40

(2) INFORMATION FOR SEQ ID NO:113:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 65 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..65

(D) OTHER INFORMATION: / Ceres Seq. ID 10221/3

(1) SEQUENCE DESCRIPTION: SEQ ID NO:113:

50 Glu Ser Leu Phe Phe Pro Ser Ala Ile Phe Phe Leu Gly Cys Arg 5  
 1 Val Lys Gly Glu Gly Phe Glu Leu Gly Phe Cys Arg Glu Lys Asp Glu 10  
 20 Pro Lys Phe Gly Asn Thr Gly Glu Ala Ser Ser Arg Gly Leu Arg Ser 25  
 35 Tyr Arg Asp Gly Gly Ala Lys Lys Glu Arg Arg Ala Leu Glu Arg Lys Tyr 30  
 40 Asp 45  
 50

(2) INFORMATION FOR SEQ ID NO:114:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 706 base pairs

378

(B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (11) MOLECULE TYPE: DNA (genomic)  
 (1x) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..706

(D) OTHER INFORMATION: / Ceres Seq. ID 1022554

(1) SEQUENCE DESCRIPTION: SEQ ID NO:114:

10 agacacaaa accgttgggt gtaagatccc aaatccacag attcccaaat aaagtaata 60  
 ctctctctct tctcaactct caccagtcac cagcagatca tcggagatgg gggagacgg 120  
 aaaaatttc accgttctcg aggtctctca gcaacatcag gccacagatt gtggagcgg- 180  
 caacgacgac aaggtttatg atgtgacaaa gttcttgat gatactcttg gtggatgag 240  
 ggttatcttg acttctacag ggaagatgc gacagatgat ttgagatg ttggacatag 300  
 ttccactcgg aaagccatgc tagatagcta ctatcggtt gatattgaca cagctactg- 360  
 gccggttaaa gtaagtttg tgcctctac gtgcagcaga gccgtgtgcta ctacggatga 420  
 gagctcgat ttgttatta agtctctca gttcttggt ccaattctaa ttatagatt- 480  
 ggctttggc attcgtaact acataagac caaggtctct tcttttga agattgaat- 540  
 gtgacctaa gtaattgctt gtatcagac tggttaact ctgtctctac ctctgtttg- 600  
 ctgttaaac accattatat tctctatgc aatcaactga tagacctttg attttaagc 660  
 cactctact tctgtctta tacaatcaa accaattac aataac

(2) INFORMATION FOR SEQ ID NO:115:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 140 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..140

(D) OTHER INFORMATION: / Ceres Seq. ID 1022555

(1) SEQUENCE DESCRIPTION: SEQ ID NO:115:

35 Met Gly Gly Asp Gly Lys Val Phe Thr Leu Ser Glu Val Ser Gln His 1  
 Ser Ser Ala Lys Asp Cys Trp Ile Val Ile Asp Gly Lys Val Tyr Asp 10  
 Val Thr Lys Phe Leu Asp Asp His Pro Gly Gly Asp Glu Val Ile Leu 15  
 Thr Ser Thr Gly Lys Asp Ala Thr Asp Phe Glu Asp Val Gly His 20  
 Ser Ser Thr Ala Lys Ala Met Leu Asp Glu Tyr Tyr Val Gly Asp Ile 25  
 40 Asp Thr Ala Thr Val Pro Val Lys Ala Lys Phe Val Pro Pro Thr Ser 30  
 Thr Lys Ala Val Ala Thr Gln Asp Lys Ser Ser Asp Phe Val Ile Lys 35  
 Leu Leu Gln Phe Leu Val Pro Leu Leu Ile Leu Gly Leu Ala Phe Gly 40  
 Ile Arg Tyr Tyr Thr Lys Thr Lys Ala Pro Ser Ser 45  
 130 135 140

(2) INFORMATION FOR SEQ ID NO:116:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 70 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..70

(D) OTHER INFORMATION: / Ceres Seq. ID 1022556



379

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:  
 Met Leu Asp Glu Tyr Tyr Val Gly Asp Ile Asp Thr Ala Thr Val Pro  
 1 5 10 15  
 Val Lys Ala Lys Phe Val Pro Thr Thr Lys Ala Val Ala Thr  
 20 25 30  
 Gln Asp Lys Ser Ser Asp Phe Val Ile Lys Leu Gln Phe Leu Val  
 35 40 45  
 Pro Leu Leu Ile Leu Gly Leu Ala Phe Gly Ile Arg Tyr Tyr Lys  
 50 55 60  
 Thr Lys Ala Pro Ser Ser  
 65 70

(2) INFORMATION FOR SEQ ID NO:117:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 722 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..722

(D) OTHER INFORMATION: / Ceres Seq. ID 1022594

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

aactcatca tctgactt tctcaagaa gagaatcaa atcannaa gt catcgagcgt  
 60  
 atcacagaa agaagaaac catggcgaa gaagaaaaa gcgtgatgaa gcaagtgatg  
 120  
 gtggcattg atgaagcga atglaagaa cgcgtcttc atagacgtt cgtgatctc  
 180  
 aaagatgac tcccgattc cgaatcatc cttctaatg ctaacatca tcttgatctc  
 240  
 agctgctt atgcctctc ttaaggccc gtcctgattg agcttaaaa cttactgcaa  
 300  
 gagatcata agaagctgg aluagatcgt cttgatgaag gaacaaaat ttgtgtgaa  
 360  
 actgggtta ctcaagaaa ggtgttgaa ttggaaac ctaagaagc gatattgaa  
 420  
 gctctgaga agctgggtt gtagctgtt gtagtgaa gcaatgtaa agagacata  
 480  
 caaagactt tcttgagg gttagcaat tactgtgta acaatgtaa ggtccagtt  
 540  
 cttgtgta gaacaaagc tgaagacct cttctatga agagttgca caaccctggc  
 600  
 ttgattgtg ttgtttgaa acattatgat aataatgct tctgtgta ctattgtt  
 660  
 aagaacctt gtgagtaaac aaattatt ttgttgata atgaagagc attgcggt  
 720  
 tc

(2) INFORMATION FOR SEQ ID NO:118:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 187 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..187

(D) OTHER INFORMATION: / Ceres Seq. ID 1022595

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:

Asn Val Ile Ile Leu Thr Phe Leu Lys Lys Glu Lys Ser Asn His Gln  
 1 5 10 15  
 Val Ile Gly Ala Ile Thr Glu Lys Lys Lys Thr Met Ala Glu Glu  
 20 25 30  
 Lys Ser Val Met Lys Gln Val Met Val Ala Ile Asp Glu Ser Glu Cys  
 35 40 45  
 Ser Lys Arg Ala Leu Gln Trp Thr Leu Val Tyr Leu Lys Asp Ser Leu  
 50 55 60  
 Ala Asp Ser Asp Ile Ile Leu Phe Thr Ala Gln Pro His Leu Asp Leu  
 65 70 75 80  
 Ser Cys Val Tyr Ala Ser Ser Tyr Gly Ala Ala Pro Ile Glu Leu Ile  
 85 90 95  
 Asn Ser Leu Gln Glu Ser His -ys Asn Ala Gly Leu Asn Arg Leu Asp  
 100 105 110

380

Glu Gly Thr Lys Ile Cys Ala Glu Thr Gly Val Thr Pro Arg Lys Val  
 115 120 125  
 Leu Glu Phe Gly Asn Pro Lys Glu Ala Ile Cys Glu Ala Ala Glu Lys  
 130 135 140  
 Leu Gly Val Asp Met Leu Val Val Gly Ser His Gly Lys Gly Ala Leu  
 145 150 155 160  
 Gln Arg Thr Phe Leu Gly Ser Val Ser Asn Tyr Cys Val Asn Asn Ala  
 165 170 175  
 Lys Cys Pro Val Leu Val Val Arg Thr Lys Ala  
 180 185

(2) INFORMATION FOR SEQ ID NO:119:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 160 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..160

(D) OTHER INFORMATION: / Ceres Seq. ID 1022596

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

Met Ala Glu Glu Glu Lys Ser Val Met Lys Gln Val Met Val Ala Ile  
 1 5 10 15  
 Asp Glu Ser Glu Cys Ser Lys Arg Ala Leu Gln Trp Thr Leu Val Tyr  
 20 25 30  
 Leu Lys Asp Ser Leu Ala Asp Ser Asp Ile Ile Leu Phe Thr Ala Gln  
 35 40 45  
 Pro His Leu Asp Leu Ser Cys Val Tyr Ala Ser Ser Tyr Gly Ala Ala  
 50 55 60  
 Pro Ile Glu Leu Ile Asn Ser Leu Gln Glu Ser His Lys Asn Ala Gly  
 65 70 75 80  
 Leu Asn Arg Leu Asp Glu Gly Thr Lys Ile Cys Ala Glu Thr Gly Val  
 85 90 95  
 Thr Pro Arg Lys Val Leu Glu Phe Gly Asn Pro Lys Glu Ala Ile Cys  
 100 105 110  
 Glu Ala Ala Glu Lys Leu Gly Val Asp Met Leu Val Val Gly Ser His  
 115 120 125  
 Gly Lys Gly Ala Leu Gln Arg Thr Phe Leu Gly Ser Val Ser Asn Tyr  
 130 135 140  
 Cys Val Asn Asn Ala Lys Cys Pro Val Leu Val Val Arg Thr Lys Ala  
 145 150 155 160

(2) INFORMATION FOR SEQ ID NO:120:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 152 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..152

(D) OTHER INFORMATION: / Ceres Seq. ID 1022597

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:

Met Lys Gln Val Met Val Ala Ile Asp Glu Ser Glu Cys Ser Lys Arg  
 1 5 10 15  
 Ala Leu Gln Trp Thr Leu Val Tyr Leu Lys Asp Ser Leu Ala Asp Ser  
 20 25 30  
 Asp Ile Ile Leu Phe Thr Ala Gln Pro His Leu Asp Leu Ser Cys Val  
 35 40 45  
 Tyr Ala Ser Ser Tyr Gly Ala Ala Pro Ile Glu Leu Ile Asn Ser Leu  
 50 55 60

381

5 Gln Glu Ser His Lys Asn Ala Gly Leu Asn Arg Leu Asp Glu Gly Thr  
65 70 75 80  
Lys Ile Cys Ala Glu Thr Gly Val Thr Pro Arg Lys Val Leu Gly Phe  
85 90 95  
Gly Asn Pro Lys Glu Ala Ile Cys Glu Ala Glu Lys Leu Gly Val  
100 105 110  
Asp Met Leu Val Val Gly Ser His Gly Lys Gly Ala Leu Glu Arg Thr  
115 120 125  
Phe Leu Gly Ser Val Ser Asn Tyr Cys Val Asn Asn Ala Lys Cys Pro  
130 135 140  
Val Leu Val Val Arg Thr Lys Ala  
145 150

(2) INFORMATION FOR SEQ ID NO:121:  
(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 610 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ix) FEATURE:  
(i) MOLECULE TYPE: DNA (genomic)

(A) NAME/KEY: -

(B) LOCATION: 1..610

(D) OTHER INFORMATION: / Ceres Seq. ID 1022621

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:

25 aatttgggt ggaattaac tggtagaga gtcaacatc caattttt ctctctctt  
atctttatc tctcaactc ttaaatggt tctttatc cgaatctctg cctctctct  
accgttgtt cagcgggta attgacctg gcataaat agcgcgttag gtacgggtg  
tgcagaagcc cagcagaca ccagactgg tgattctg agcagcgag ctatagccg  
tggtttagg tgcagccgg ttattggat gctactctac agcttgaaga cccagggatg  
tggcttacct ccgcagccg ccgatttat cggagcgttg gaagtgta gtcacttgg  
gagtttggc atcgagggt gtttttga cactaaaca aaactgggt cagttctgc  
aaatgggca ttgtttgt tgggtcgtt cgaagttta tgcattctt cgtctctag  
cattctggg gttttgta ttcattctt ggtataggg tggttccag tgcactctc  
tggagccgg gttttgtt aaccaggtt agttcgtat aataaagtg tcaattaat  
aattgtctc

(2) INFORMATION FOR SEQ ID NO:122:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 158 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) FEATURE:  
(i) MOLECULE TYPE: peptide

(A) NAME/KEY: peptide

(B) LOCATION: 1..158

(D) OTHER INFORMATION: / Ceres Seq. ID 1022622

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

50 Met Leu Leu Leu Ser Pro Ile Ser Ala Ser Leu Pro Pro Ser Phe His  
1 5 10 15  
Arg Gly Asn Leu Ile Arg Arg Ser Ile Lys Pro Leu Gly Arg Val Val  
20 25 30  
Ala Lys Ala Lys Asp Asn Thr Asp Thr Gly Phe Leu Glu Thr Ala  
35 40 45  
Ala Ile Ala Gly Gly Leu Val Ser Thr Pro Val Ile Gly Trp Ser Leu  
50 55 60  
Tyr Thr Leu Lys Thr Thr Gly Cys Gly Leu Pro Pro Gly Pro Ala Gly  
65 70 75 80  
Leu Ile Gly Ala Leu Glu Gly Val Ser Tyr Leu Val Val Gly Ile  
85 90 95  
Val Gly Trp Ser Leu Tyr Thr Lys Thr Cys Thr Gly Ser Gly Leu Pro  
100 105 110  
Asn Gly Pro Phe Gly Leu Leu Gly Ala Val Glu Gly Leu Ser Tyr Leu

382

5 Ser Val Leu Ala Ile Leu Val Val Phe Gly Ile Gln Phe Leu Asp Asn  
115 120 125  
Gly Ser Val Pro Gly Pro Leu Pro Ser Asp Gln Cys Phe Gly  
130 135 140  
145 150 155

(2) INFORMATION FOR SEQ ID NO:123:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 872 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ix) FEATURE:  
(i) MOLECULE TYPE: DNA (genomic)

(A) NAME/KEY: -

(B) LOCATION: 1..872

(D) OTHER INFORMATION: / Ceres Seq. ID 1024375

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:

20 atcgacatt aaaaactgat ttactgaaa ctgtgtcgc aaggagaac cgcgaacct  
littgtcaa tctagggttc tctctgac ttgagatt cgtctcgtt tctactaga  
tcgtgaga tgtctccgc tgaagacag atcaataag acaagatgc tgaagcaaa  
gaatgtagg agcaagtgc tgaagcttg ttgatttgc aaaaactaa cctgagttg  
aaaagcagt tgaagatct ctacatcac caagctgtc acatgatat cctcgaaac  
cgcaagctg tctgattta cgttcacatc agattgaga aagcttccg caagatat  
ccctctcgc tcaagaact tgaagagc ttcagtgga agatgttat cttgttacc  
25 acaagaga tcaagctgc cccaagaag ggtgtcgtc ttcagagcc acgtaacga  
accttacct cagttcatg agctatggt gaagatgtg ctttccgc tgaatgtt  
ggaagcgtt ctcctacgc tctatggt tccaaatca tgaagcttt ttggtatgc  
aagaaaga acacacaga gtacagctc gacatactg tgggtgta ccytaaat  
acgtgcaag agtctgtt tgaatcaca gtcagagctt gaaagagat gatgagaa  
30 catcgata gtgaagaga gctttgtt agtttgtg gtattaga tgaagaaac  
tctctgatt cagttcgtt ttcacatct ttaagtctt attacaatg actactttt  
gtttttca atttcaaac cttatgatt tc

(2) INFORMATION FOR SEQ ID NO:124:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 190 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) FEATURE:  
(i) MOLECULE TYPE: peptide

(A) NAME/KEY: peptide

(B) LOCATION: 1..190

(D) OTHER INFORMATION: / Ceres Seq. ID 1024376

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:124:

45 Met Phe Ser Ala Gln Asn Lys Ile Asn Lys Asp Lys Asn Ala Glu Pro  
1 5 10 15  
Thr Glu Cys Glu Glu Gln Val Ala Gln Ala Leu Phe Asp Leu Glu Asn  
20 25 30  
Thr Asn Gln Glu Leu Lys Ser Glu Leu Lys Asp Leu Tyr Ile Asn Gln  
35 40 45  
Ala Val His Met Asp Ile Ser Gly Asn Arg Lys Ala Val Val Ile Tyr  
50 55 60  
Val Pro Phe Arg Leu Arg Lys Ala Phe Arg Lys Ile His Pro Arg Leu  
65 70 75 80  
Val Arg Glu Leu Glu Lys Lys Phe Ser Gly Lys Asp Val Ile Phe Val  
85 90 95  
Thr Thr Arg Arg Ile Met Arg Pro Pro Lys Lys Gly Ala Ala Val Gln  
100 105 110  
Arg Pro Arg Asn Arg Thr Leu Thr Ser Val His Glu Ala Met Leu Glu  
115 120 125  
Asp Val Ala Phe Pro Ala Glu Ile Val Gly Lys Arg Thr Arg Tyr Arg  
130 135 140

384

Leu Asp Gly Ser Lys Ile Met Lys Val Phe Leu Asp Ala Lys Glu Lys  
145 150 155 160  
Asn Asn Thr Glu Tyr Lys Leu Glu Thr Met Val Gly Val Tyr Arg Lys  
165 170 175  
5 Leu Thr Gly Lys Asp Val Phe Glu Tyr Pro Val Glu Ala  
180 185 190

(2) INFORMATION FOR SEQ ID NO:125:  
(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 139 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..139

(D) OTHER INFORMATION: / Ceres Seq. ID 1024377

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:125:

Met Asp Ile Ser Gly Asn Arg Lys Ala Val Val Ile Tyr Val Pro Phe  
1 5 10 15

Arg Leu Arg Lys Ala Phe Arg Lys Ile His Pro Arg Leu Val Arg Glu  
20 25 30

Leu Glu Lys Lys Phe Ser Gly Lys Asp Val Ile Phe Val Thr Thr Arg  
35 40 45

Arg Ile Met Arg Pro Pro Lys Lys Gly Ala Val Gln Arg Pro Arg  
50 55 60

Asn Arg Thr Leu Thr Ser Val His Glu Ala Met Leu Glu Asp Val Ala  
65 70 75 80

Phe Pro Ala Glu Ile Val Gly Lys Arg Thr Arg Tyr Arg Leu Asp Gly  
85 90 95

Ser Lys Ile Met Lys Val Phe Leu Asp Ala Lys Glu Lys Asn Asn Thr  
100 105 110

Glu Tyr Lys Leu Glu Thr Met Val Gly Val Tyr Arg Lys Leu Thr Gly  
115 120 125

Lys Asp Val Val Phe Glu Tyr Pro Val Glu Ala  
130

(2) INFORMATION FOR SEQ ID NO:126:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 545 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..545

(D) OTHER INFORMATION: / Ceres Seq. ID 1024535

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

aaaaattgtt aagttctgtt cccggagacc agagttgttg tccggttata cgtattctgc 60  
g-accaaac ccccgagaa gttgttcagg tgggaagga tcttctgcta cagatgtccc 120  
ggcaggaag agccaccgtt gaggagacca ctgattctga ttggttgag tctcgacgag 180  
agaattcttg tcaagctcgg gtcagaccgc gaacttcgg gcaagttcca cgcgtttgat 240  
cagatttga atagattctt ggggtgattt gaagaaacta tcaatcagat agaaatcgat 300  
gacgaacat atgaagagat tttctgacctt acaagacgga cagattgatt tctattctgt 360  
agaggagat gattgattt ggtgtctcca cgcgtgagga cagagctg agttcaact 420  
caaatcttt gttctcttt tctgaaggg gaatttgga ctattctgt attagttcac 480  
ttctgtctt aactaaaca agctgaaga agtagtttg cttaaatctt caatgcttt 540  
cagac

(2) INFORMATION FOR SEQ ID NO:127:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 99 amino acids

(B) TYPE: amino acid

384

(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..98

(D) OTHER INFORMATION: / Ceres Seq. ID 1024536

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:

Met Ser Gly Glu Glu Glu Ala Thr Val Arg Glu Pro Leu Asp Leu Ile  
1 5 10 15

Arg Leu Ser Leu Asp Glu Arg Ile Tyr Val Lys Leu Arg Ser Arg  
20 25 30

Glu Leu Arg Gly Lys Leu His Ala Phe Asp Gln His Leu Asn Met Ile  
35 40 45

Leu Gly Asp Val Glu Glu Thr Ile Thr Thr Val Glu Ile Asp Asp Glu  
50 55 60

Thr Tyr Glu Glu Ile Val Arg Thr Thr Lys Arg Thr Ile Glu Phe Leu  
65 70 75 80

Phe Val Arg Gly Asp Gly Val Ile Leu Val Ser Pro Leu Arg Thr  
85 90 95

Ala Ala

(2) INFORMATION FOR SEQ ID NO:128:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 53 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..53

(D) OTHER INFORMATION: / Ceres Seq. ID 1024537

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

Lys Leu Leu Ser Ser Cys Pro Arg Asp Gln Ser Leu Cys Pro Val Ile  
1 5 10 15

Arg Phe Val Gly Thr Lys Thr Leu Arg Glu Val Phe Gln Val Gly Arg  
20 25 30

Tyr Leu Cys Leu His Asp Val Arg Arg Gly Arg Ser His Arg Glu Gly  
35 40 45

Ala Thr Arg Ser Asp  
50

(2) INFORMATION FOR SEQ ID NO:129:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 52 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..52

(D) OTHER INFORMATION: / Ceres Seq. ID 1024538

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

Met Ile Leu Gly Asp Val Glu Glu Thr Ile Thr Thr Val Glu Ile Asp  
1 5 10 15 25

Asp Glu Thr Tyr Glu Glu Ile Val Arg Thr Thr Lys Arg Thr Ile Glu  
20 25 30

Phe Leu Phe Val Arg Gly Asp Gly Val Ile Leu Val Ser Pro Leu  
35 40 45

Arg Thr Ala Ala  
50

385

## (2) INFORMATION FOR SEQ ID NO:130:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 797 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..797

(D) OTHER INFORMATION: / Ceres Seq. ID 1025683

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

aacctttaa cccgttacc ttaacctca agccctgcgc tcaaatgac cccctcagag  
 acgaaaaat gaagaccat ttctctcgg agctactaga catccccac ggcctgcga  
 ttaagttaa cgcacagatg attgagtcg aggtccacg aggtcaact actctgact  
 tcaagcatct gaattcgat ttccagt-ga ttaagacca agtcaatga aacgtcagc  
 ttaagatga ticttggtt ggtctcga agacaatgc ttcataga actgcttaa  
 gcatatga taatctcatt gctgggtta ccaaggtt tctctaga atgagattg  
 tgcattgca tttctcatt actctctc l-galgtaa caaagagat attgagattc  
 gtaactct tggtagaag aggtgagga aggttgagat gttgaggt gttaaattg  
 ltcgaactga gaagtagaag gagagatta tcttgaggg aaatcatatt gacttgatt  
 cagggcttg tctttgatc aa-cagaaat gtcattgaa gaagagagat atcaggagc  
 tcttgaggg taccatag agggagaag gcaagatcgc agtcaggaa tgaattcgc  
 tattgaagt tctatagg atttatag gtgaagcag gatattatcg tagcttttg  
 tttacattc tcttatgga ttgtgattt tgtttctc ttgaaatc tgaatgacg  
 aaattcatta tccgttc

## (2) INFORMATION FOR SEQ ID NO:131:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 216 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..216

(D) OTHER INFORMATION: / Ceres Seq. ID 1025684

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

Pro Leu Asn Ser Val Thr Leu Asn Leu Lys Pro Ser Ser Pro Lys Phe  
 1 5  
 Pro Leu Arg Asp Glu Lys Met Lys Thr Ile Leu Ser Ser Glu Thr Met  
 20 25  
 Asp Ile Pro Asp Gly Val Ala Ile Lys Val Asn Ala Lys Val Ile Glu  
 35 40 45  
 Val Glu Gly Pro Arg Gly Lys Leu Thr Arg Asp Phe Lys His Leu Asn  
 50 55 60  
 Leu Asp Phe Glu Leu Ile Lys Asp Glu Val Thr Gly Lys Arg Glu Leu  
 65 70 75  
 Lys Ile Asp Ser Trp Phe Gly Ser Arg Lys Thr Ser Ala Ser Ile Arg  
 80 85 90  
 Thr Ala Leu Ser His Val Asp Asn Leu Ile Ala Gly Val Thr Glu Gly  
 100 105 110  
 Phe Leu Tyr Arg Met Arg Phe Val Tyr Ala His Phe Pro Ile Asn Ala  
 115 120 125  
 Ser Ile Asp Gly Asn Asn Lys Ser Ile Glu Ile Arg Asn Phe Leu Gly  
 130 135 140  
 Glu Lys Lys Val Arg Lys Val Glu Met Leu Asp Gly Val Lys Ile Val  
 145 150 155  
 Arg Ser Glu Lys Val Lys Asp Glu Ile Ile Leu Glu Gly Asn Asp Ile  
 160 165 170  
 Glu Leu Val Ser Arg Ser Cys Ala Leu Ile Asn Glu Lys Cys His Val  
 175 180 185 190

386

Lys Lys Lys Asp Ile Arg Lys Phe Leu Asp Gly Ile Tyr Val Ser Glu  
 195 200  
 Lys Gly Lys Ile Ala Val Glu Glu  
 210 215

## (2) INFORMATION FOR SEQ ID NO:132:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 194 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..194

(D) OTHER INFORMATION: / Ceres Seq. ID 1025685

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

Met Lys Thr Ile Leu Ser Ser Glu Thr Met Asp Ile Pro Asp Gly Val  
 1 5 10  
 Ala Ile Lys Val Asn Ala Lys Val Ile Glu Val Glu Gly Pro Arg Gly  
 20 25 30  
 Lys Leu Thr Arg Asp Phe Lys His Leu Asn Leu Asp Phe Glu Leu Ile  
 35 40 45  
 Lys Asp Glu Val Thr Gly Lys Arg Glu Leu Lys Ile Asp Ser Trp Phe  
 50 55 60  
 Gly Ser Arg Lys Thr Ser Ala Ser Ile Arg Thr Ala Leu Ser His Val  
 65 70 75 80  
 Asp Asn Leu Ile Ala Gly Val Thr Glu Gly Phe Leu Tyr Arg Met Arg  
 85 90  
 Phe Val Tyr Ala His Phe Pro Ile Asn Ala Ser Ile Asp Gly Asn Asn  
 100 105 110  
 Lys Ser Ile Glu Ile Arg Asn Phe Leu Gly Glu Lys Val Arg Lys  
 115 120 125  
 Val Glu Met Leu Asp Gly Val Lys Ile Val Arg Ser Glu Lys Val Lys  
 130 135 140  
 Asp Glu Ile Ile Leu Glu Gly Asn Asp Ile Glu Leu Val Ser Arg Ser  
 145 150 155  
 Cys Ala Leu Ile Asn Glu Lys Cys His Val Lys Lys Asp Ile Arg  
 160 165 170 175  
 Lys Phe Leu Asp Gly Ile Tyr Val Ser Glu Lys Gly Lys Ile Ala Val  
 180 185 190  
 Glu Glu

## (2) INFORMATION FOR SEQ ID NO:133:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 185 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..185

(D) OTHER INFORMATION: / Ceres Seq. ID 1025666

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

Met Asp Ile Pro Asp Gly Val Ala Ile Lys Val Asn Ala Lys Val Ile  
 1 5 10  
 Glu Val Glu Gly Pro Arg Gly Lys Leu Thr Arg Asp Phe Lys His Leu  
 20 25 30  
 Asn Leu Asp Phe Glu Leu Ile Lys Asp Glu Val Thr Gly Lys Arg Glu  
 35 40 45  
 Leu Lys Ile Asp Ser Trp Phe Gly Ser Arg Lys Thr Ser Ala Ser Ile  
 50 55 60

387  
 Arg Thr Ala Leu Ser His Val Asp Asn Leu Ile Ala Gly Val Thr Gln 70 80  
 Gly Phe Leu Tyr Arg Met Arg Phe Val Tyr Ala His Phe Pro Ile Asn 75 85  
 Ala Ser Ile Asp Gly Asn Asn Lys Ser Ile Glu Ile Arg Asn Phe Leu 100 110  
 Gly Glu Lys Val Arg Lys Val Glu Met Leu Asp Gly Val Lys Ile 105 115  
 Val Arg Ser Glu Lys Val Lys Asp Glu Ile Ile Leu Glu Gly Asn Asp 120 125  
 Ile Glu Leu Val Ser Arg Ser Cys Ala Leu Ile Asn Gln Lys Cys His 130 135  
 Val Lys Lys Lys Asp Ile Arg Lys Phe Leu Asp Gly Ile Tyr Val Ser 140 145  
 Glu Lys Gly Lys Ile Ala Val Glu Glu 150 155  
 (2) INFORMATION FOR SEQ ID NO:134:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 55 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..55  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1027152  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:  
 gtttttca tctgtctc cttctctc accgcacaga tccaagacat ttgtattaca 60  
 cttcatccc ggcgaatatg ggcagcgaaa agaaacgaa gaagtcacat gaggtatca 120  
 acagcaggtt ggtctctgtt atgagagtg gtaaacacac cttgtgtac agtcgttc 180  
 tcaatctct cgtgtctcc aaaggaacac tgttttaac cttccacaaat tgcacacgt 240  
 tgaagacat agagattag tactacgaa -gtcgttaa agtgggtgc caccattaca 300  
 aggaataa cgttgattg gaacacatt cgcgaagta cttcgtgtt ttgtgtcga 360  
 gattgtga cttcgtgat ttgcacata tcaagcaat cttcgtgac cagtgatca 420  
 attcgtat ttattttc cagacttt taccgttt ttgtatgag tctctcaaa 480  
 tttatgac aatagacat ggtgtctc ttgtcact ttcttatat tgaggaccgg 540  
 aaacaaatt tcttt

25  
 (2) INFORMATION FOR SEQ ID NO:135:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 137 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..137  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1027153  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:  
 Phe Cys His Arg Ala Pro Leu Leu Leu His Arg Thr Asp Pro Arg His 1 15  
 Phe Asp Tyr Thr Pro Ser Pro Ala Lys Met Val Thr Glu Lys Lys Thr 20 30  
 Lys Lys Ser His Glu Gly Ile Asn Ser Arg Leu Ala Leu Val Met Lys 35 45  
 Ser Gly Lys Tyr Thr Leu Gly Tyr Lys Ser Val Leu Lys Ser Leu Arg 50 60  
 Gly Ser Lys Gly Lys Leu Ile Leu Ile Ser Thr Asn Cys Pro Pro Leu 65 75  
 Arg Arg Ser Glu Ile Glu Tyr Tyr Ala Met Leu Ala Lys Val Gly Val 80 95

388  
 His His Tyr Asn Gly Asn Asn Val Asp Leu Gly Thr Ala Cys Gly Lys 100 105  
 Tyr Phe Arg Val Ser Cys Leu Ser Ile Val Asp Pro Gly Asp Ser Asp 110 115  
 Ile Ile Lys Ser Ile Pro Gly Asp Gln 120 125  
 (2) INFORMATION FOR SEQ ID NO:136:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 112 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..112  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1027154  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:136:  
 Met Val Thr Glu Lys Lys Thr Lys Lys Ser His Glu Gly Ile Asn Ser 1 10  
 Arg Leu Ala Leu Val Met Lys Ser Gly Lys Tyr Thr Leu Gly Tyr Lys 15 20  
 Ser Val Leu Lys Ser Leu Arg Gly Ser Lys Gly Lys Leu Ile Leu Ile 25 30  
 Ser Thr Asn Cys Pro Pro Leu Arg Arg Ser Glu Ile Glu Tyr Tyr Ala 35 40  
 Met Leu Ala Lys Val Gly Val His His Tyr Asn Gly Asn Asn Val Asp 45 50  
 Leu Gly Thr Ala Cys Gly Lys Tyr Phe Arg Val Ser Cys Leu Ser Ile 55 60  
 Val Asp Pro Gly Ser Asp Ile Ile Lys Ser Ile Pro Gly Asp Gln 65 70  
 (2) INFORMATION FOR SEQ ID NO:137:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 91 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..91  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1027155  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:  
 Met Lys Ser Gly Lys Tyr Thr Leu Gly Tyr Lys Ser Val Leu Lys Ser 1 10  
 Leu Arg Gly Ser Lys Gly Lys Leu Ile Leu Ile Ser Thr Asn Cys Pro 15 20  
 Pro Leu Arg Arg Ser Glu Ile Glu Tyr Tyr Ala Met Leu Ala Lys Val 25 30  
 Gly Val His His Tyr Asn Gly Asn Asn Val Asp Leu Gly Thr Ala Cys 35 40  
 Gly Lys Tyr Phe Arg Val Ser Cys Leu Ser Ile Val Asp Pro Gly Asp 45 50  
 Ser Asp Ile Ile Lys Ser Ile Pro Gly Asp Gln 55 60  
 (2) INFORMATION FOR SEQ ID NO:138:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 653 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

389

(11) MOLECULE TYPE: DNA (genomic)

(1x) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..653

(D) OTHER INFORMATION: / Ceres Seq. ID 1028095

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:138:

5 aataaagtt ttattatc tatctatcca attgagacg aatcatatcg agcatgatg 60  
 ttggagac ttggcgat ttgcagacc gattctccg gcgcgtttt cgaacccgg 120  
 tgggtttt ggtgcgac gcgtctgc agttccatc agtctccct cgtccctac 180  
 attccgga tattgcttc ttgcagatc ttgatctta agtctgctt aaagaagac 240  
 attgattact cctctacga cgaagcgag tggagattga agtctgctt ttctatagc 300  
 taatcgag cattgttct ctatgtgct tctgttgct tggatgaca agattcggt 360  
 tgaagactg ggcctaac ttgactggt gggctgggt tcttcaatg tcatctgta 420  
 ttataagt gactaaga ttgacatcg cactcaagt agcgacagc cacatactc 480  
 attccgact cagattcaca ctgacacac ttacaagttc ctctgttct tctactagt 540  
 aaattcgtt gctgtgctt ctacattcg cattctgtc gacatcttc tttgcatca 600  
 ggtggttt cgtcaaac ttgtgtgaa tgttatgaa aatactccaa atc

(2) INFORMATION FOR SEQ ID NO:138:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 135 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..135

(D) OTHER INFORMATION: / Ceres Seq. ID 1028096

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:139:

30 Met Asp Leu Ala Glu Leu Trp Ala Ile Phe Gly Pro Gly Phe Ser Gly 1  
 1 Ala Val Phe Gly Thr Gly Trp Thr Phe Thr Val Asp Ala Val Val Cys 15  
 20 Ser Ile Gln Val Pro Phe Val His Tyr Leu Pro Gly Ile Phe Ala 30  
 35 Ser Leu Gly Ala Leu Met Phe Asn Cys Val Arg Lys Glu Asp Ile Asp 45  
 50 Tyr Ser Pro Tyr Asp Glu Gly Glu Trp Arg Leu Lys Leu Trp Leu Phe 60  
 65 Ile Ala Tyr Val Val Ala Phe Val Ser Leu Ala Ala Ser Val Gly Leu 75  
 80 Leu Ile Gln Asp Ser Val Val Lys Thr Gly Pro Ser Thr Trp Thr Gly 95  
 100 Val Ala Gly Val Phe Gln Cys Val Phe Val Leu Ile Ser Gly Leu Met 110  
 115 Tyr Trp Thr Ser His Ser Glu 120  
 130

(2) INFORMATION FOR SEQ ID NO:140:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 82 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..82

(D) OTHER INFORMATION: / Ceres Seq. ID 1028097

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:140:

60 Met: Phe Asn Cys Val Arg Lys Glu Asp Ile Asp Tyr Ser Pro Tyr Asp 1  
 5 Gln Gly Glu Trp Arg Leu Lys Leu Trp Leu Phe Ile Ala Tyr Val Val 15

390

(11) MOLECULE TYPE: DNA (genomic)

(1x) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..779

(D) OTHER INFORMATION: / Ceres Seq. ID 1028098

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:141:

5 Ala Phe Val Ser Leu Ala Ala Ser Val Gly Leu Leu Ile Gln Asp Ser 30  
 35 Val Val Lys Thr Gly Pro Ser Thr Trp Thr Gly Val Ala Gly Val Phe 45  
 50 Gln Cys Val Phe Val Leu Ile Ser Gly Leu Met Tyr Trp Thr Ser His 60  
 65 Ser Glu 75

(2) INFORMATION FOR SEQ ID NO:141:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 80 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..80

(D) OTHER INFORMATION: / Ceres Seq. ID 1028098

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:141:

10 Ile Lys Phe Leu Phe Ile Tyr Leu Ser Arg Trp Arg Arg Ile Ile Ser 15  
 1 Ser Asp Gly Phe Gly Arg Thr Val Gly Asp Phe Arg Thr Arg Ile Leu 30  
 20 Arg Arg Arg Phe Arg Asn Arg Val Val Leu Gly Arg Arg Arg Arg 35  
 40 Leu Gln Phe His Pro Ser Ser Leu Arg Pro Leu Pro Ser Arg His Ile 45  
 50 Arg Phe Ser Arg Ser Phe Asp Val Gln Leu Arg Gln Lys Arg Arg His 60  
 65

(2) INFORMATION FOR SEQ ID NO:142:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 779 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: DNA (genomic)

(1x) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..779

(D) OTHER INFORMATION: / Ceres Seq. ID 1028098

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:142:

45 gctcaaacct taaactctaa agactttgic attctgtgaa cgtctgtcat aactgtact 60  
 tccaaagaat tctcaaaatc cctaaacttc gaaatggca gaaaccccg agaatctac 120  
 caccacgact tcatcgccg caccggatc cgaagctcc accactatt ctgcactgac 180  
 agatccaca tctcaagaag cagcttcaaa agacacagat ctgacgaag agccgaatc 240  
 agagaagaa ccaggagaaa tctctctcg aatctggca ccgacacga aactcgcga 300  
 cccgtctcg aatcgctga tggagacctt accaccgaa tcaatctct taagagata 360  
 cggactctt aaatctgag atgcaaccac cgtcgcaaa ctatctgaag aagagctta 420  
 tgggttgtt tgaatgttg tgcgagta tgaatgttg gctatgttg tgaatgttg 480  
 tttcaaaag attagtaaa gtagcttaa atctgttaa gctagatcta atgtgtgtg 540  
 tggaaatgga agtctgggg agtcaatac tgaatgagt gaggttcta aagatgtgc 600  
 tggctcgtt tcaagagag agaaagta ggtctgaga attgtgtt acttctcat 660  
 taaatcgt tgtcttag ggtttctg tcttgaat gatgatgat actgtgtgat 720  
 ttgttatat gcaattga agatgttct tctcttaa tagagtgca agtgtcttc

(2) INFORMATION FOR SEQ ID NO:143:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 180 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

391

(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide  
(B) LOCATION: 1..180

(D) OTHER INFORMATION: / Ceres Seq. ID 1028609

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:

Met Ala Glu Thr Ala Glu Thr Ile Asn Thr Thr Ile Ser Ser Pro Pro

1 5 10 15

Pro Glu Ser Glu Ser Thr Thr Ile Ser Ala Met Thr Asp Pro Thr

10 20 25 30

Ser Gln Glu Ala Ala Ser Lys Asp Thr Asp Leu Thr Lys Glu Ala Glu

35 40 45

Ser Glu Lys Lys Pro Gly Ile Ser Leu Arg Ile Trp Pro Pro Thr

50 55 60

Gln Lys Thr Arg Asp Ala Val Leu Asn Arg Leu Ile Glu Thr Leu Ser

65 70 75 80

Thr Glu Ser Ile Leu Ser Lys Arg Tyr Gly Thr Leu Lys Ser Asp Asp

85 90 95

Ala Thr Thr Val Ala Lys Leu Ile Glu Glu Ala Tyr Gly Val Ala

100 105 110

Ser Asn Ala Val Ser Ser Asp Asp Gly Ile Lys Ile Leu Glu Leu

115 120 125

Tyr Ser Lys Glu Ile Ser Lys Arg Met Leu Glu Ser Val Lys Ala Arg

130 135 140

Ser Asn Ala Ser Val Gly Asn Gly Ser Val Glu Asp Ala Asn Thr Asp

145 150 155

Ala Ser Glu Val Ser Lys Asp Ala Gly Pro Gly Ser Glu Glu Glu

160 165 170 175

Lys Ser Glu Ala

180

(2) INFORMATION FOR SEQ ID NO:144:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 153 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..153

(D) OTHER INFORMATION: / Ceres Seq. ID 1028610

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:

Met Thr Asp Pro Thr Ser Gln Glu Ala Ala Ser Lys Asp Thr Asp Leu

1 5 10 15

Lys Glu Ala Glu Ser Glu Lys Lys Pro Gly Ile Ser Leu Arg

20 25 30

Ile Trp Pro Pro Thr Gln Lys Thr Arg Asp Ala Val Leu Asn Arg Leu

35 40 45

Ile Glu Thr Leu Ser Thr Glu Ser Ile Leu Ser Lys Arg Tyr Gly Thr

50 55 60

Leu Lys Ser Asp Ala Thr Thr Val Ala Lys Leu Ile Glu Glu Glu

65 70 75 80

Ala Tyr Gly Val Ala Ser Asn Ala Val Ser Ser Asp Asp Gly Ile

85 90 95

Lys Ile Leu Glu Leu Tyr Ser Lys Glu Ile Ser Lys Arg Met Leu Glu

100 105 110

Ser Val Lys Ala Arg Ser Asn Ala Ser Val Gly Asn Gly Ser Val Glu

115 120 125

Asp Ala Asn Thr Asp Ala Ser Glu Val Ser Lys Asp Asp Ala Gly Pro

130 135 140

Gly Ser Glu Glu Lys Ser Glu Ala

145 150

392

(2) INFORMATION FOR SEQ ID NO:145:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 94 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..94

(D) OTHER INFORMATION: / Ceres Seq. ID 1028611

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:

Met Gln Pro Pro Arg Asn Leu Ser Lys Arg Leu Met Val Leu

1 5 10 15

Leu Arg Met Leu Cys Arg Val Met Met Gly Leu Arg Phe Leu Ser

20 25 30

Phe Ile Leu Lys Arg Leu Val Arg Gly Cys Leu Asn Leu Leu Arg Leu

35 40 45

Asp Leu Met Leu Val Leu Glu Met Glu Val Trp Arg Met Leu Ile Leu

50 55 60

Met Leu Val Arg Phe Leu Lys Met Met Leu Val Leu Val Gln Arg Arg

65 70 75 80

Arg Arg Val Arg Leu Glu Lys Phe Gly Phe Thr Ser Leu Leu

85 90

(2) INFORMATION FOR SEQ ID NO:146:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 673 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..673

(D) OTHER INFORMATION: / Ceres Seq. ID 1030069

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:146:

atccaaataa aacaaataca aaattatat tcaagagaaa aggaataaaa tgaattcat

60

ctccgacag gtaagaac tctcagctt aacacagag gacacagcc acacaagcc

120

agtcgaagg accgaacag ctcaagcgc agctaccac gccagctca tggcaagtc

180

gggtaaagtc gccgagcct ctgtgatat cttagacct ccgagaaat accgaagtt

240

cgatgaagag agtagactg gtacactct ccacaagct gagaagatc tcaagacta

300

cgagctcca cactcaacy gtcgtgttg tctctctct ccagagatc agctgagcc

360

aggaagtcag cc-gagcgg cggctcaaga agacgatgaa ggtcttggt ggggcttgg

420

aggtatgcc aagatggctc aagtttctt gaagttatt gactttaat tgtgttcat

480

catttcgta ataataatt aaataactag tatcgtttgt gactagtta tgtgttcgt

540

ttatgctta tgggagtaga cgaagagtg taataacttc tggatcatc gnatcatac

600

cattctgtt gtc

660

(2) INFORMATION FOR SEQ ID NO:147:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 171 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..171

(D) OTHER INFORMATION: / Ceres Seq. ID 1030070

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:147:

Ser Lys Lys Asn Lys Thr Lys Asn Tyr Ile Gln Glu Lys Lys Glu Lys

1 5 10 15

393

Met Asn Phe Ile Ser Asp Gln Val Lys Lys Leu Ser Ser Thr Pro  
20 30  
Glu Glu Pro Asp His Asn Lys Pro Val Glu Gly Thr Glu Thr Ala Thr  
35 40 45  
5 Arg Pro Ala Thr Asn Ala Glu Leu Met Ala Ser Ala Lys Val Val Ala  
50 60  
Glu Ala Ala Gln Ala Ala Arg Asn Glu Ser Asp Lys Leu Asp Lys  
65 70 75 80  
Gly Lys Val Ala Gly Ala Ser Ala Asp Ile Leu Asp Ala Ser Glu Lys  
85 90  
10 Tyr Gly Lys Phe Asp Glu Lys Ser Ser Thr Gly His Tyr Leu Asp Lys  
100 105 110  
Ala Glu Lys Tyr Leu Asn Asp Tyr Glu Ser Ser His Ser Thr Gly Ala  
115 120 125  
15 Gly Gly Pro Pro Pro Thr Ser Gln Ala Glu Pro Ala Ser Gln Pro  
130 135 140  
Glu Pro Ala Ala Lys Lys Asp Asp Glu Glu Ser Gly Gly Lys Leu Gly  
145 150 155 160  
Gly Tyr Ala Lys Met Ala Gln Gly Phe Leu Lys  
165 170

(2) INFORMATION FOR SEQ ID NO:148:  
(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 155 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) FEATURE:

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..155

(D) OTHER INFORMATION: / Ceres Seq. ID 1030071  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:148:

Met Asn Phe Ile Ser Asp Gln Val Lys Lys Leu Ser Ser Thr Pro  
1 5 10 15  
35 Glu Glu Pro Asp His Asn Lys Pro Val Glu Gly Thr Glu Thr Ala Thr  
20 25 30  
Arg Pro Ala Thr Asn Ala Glu Leu Met Ala Ser Ala Lys Val Val Ala  
35 40 45  
40 Glu Ala Ala Gln Ala Ala Arg Asn Glu Ser Asp Lys Leu Asp Lys  
50 55 60  
Gly Lys Val Ala Gly Ala Ser Ala Asp Ile Leu Asp Ala Ser Glu Lys  
65 70 75 80  
Tyr Gly Lys Phe Asp Glu Lys Ser Ser Thr Gly His Tyr Leu Asp Lys  
85 90 95  
45 Ala Glu Lys Tyr Leu Asn Asp Tyr Glu Ser Ser His Ser Thr Gly Ala  
100 105 110  
Gly Gly Pro Pro Pro Thr Ser Gln Ala Glu Pro Ala Ser Gln Pro  
115 120 125  
Glu Pro Ala Ala Lys Lys Asp Asp Glu Glu Ser Gly Gly Lys Leu Gly  
130 135 140  
145 Gly Tyr Ala Lys Met Ala Gln Gly Phe Leu Lys  
150 155

(2) INFORMATION FOR SEQ ID NO:149:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 115 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) FEATURE:

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..115

394

(D) OTHER INFORMATION: / Ceres Seq. ID 1030072  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:149:  
Met Ala Ser Ala Lys Val Val Ala Glu Ala Ala Gln Ala Ala Arg  
1 5 10 15  
5 Asn Glu Ser Asp Lys Leu Asp Lys Gly Lys Val Ala Gly Ala Ser Ala  
20 25 30  
Asp Ile Leu Asp Ala Ser Glu Lys Tyr Gly Lys Phe Asp Glu Lys Ser  
35 40 45  
10 Ser Thr Gly His Tyr Leu Asp Lys Ala Glu Lys Tyr Leu Asn Asp Tyr  
50 55 60  
Glu Ser Ser His Ser Thr Gly Ala Gly Gly Pro Pro Pro Thr Ser  
65 70 75 80  
Gln Ala Glu Pro Ala Ser Gln Pro Glu Pro Ala Ala Lys Lys Asp Asp  
85 90 95  
15 Glu Glu Ser Gly Gly Lys Leu Gly Tyr Ala Lys Met Ala Gln Gly  
100 105 110  
Phe Leu Lys  
115

(2) INFORMATION FOR SEQ ID NO:150:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 944 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ix) FEATURE:

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..944

(D) OTHER INFORMATION: / Ceres Seq. ID 1032069

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:150:

atggcgaatt tcaatggatc actctctc tctatctctg tcaaacatc tctccaaag  
60  
taggcagaag aagatgggac cgcatttga tctatcaga gcagagctag cactgtggt  
120  
aatgatttg acaaaagag agcagaagag taagtattgc agagctaac agtattgtc  
180  
caagtcttg atgtgtggac aactgtgtac tctcaaat gtgcacaa ctactagctt  
240  
agcaagaaa gtcttcgtc tttaaatg ttcaaatg ttcaatgct ttaacatcc  
300  
ttgtctaaa ggaactctc tctctctg ttacttgg aag-cgaaga agcactttt  
360  
atctaatc ttgtcttg atcaattg atgtgttgg agatcagaa tatatagaa  
420  
caagaacga gctgattac ttgacgtat atctcttc tctgtgaggg gattctctt  
480  
c-gcacaact ttgctgag ttgtgagat ggaagagctt tctctcaa tgaagaagat  
540  
aaatcaaac gaggggtcac ttgtttgat caatcagct atggacattg ttgtagcag  
600  
ttgtctctt cagtttagc caacgaagat cactctctgt g-caccgag ctttggatt  
660  
ctacactcc atcatttct g-taccagtt gcttccgaca cgcctcaaga tcaaacacc  
720  
ctgaagtaa tggaggaagc tgggtttaag ggaagaagca ttccagaag ttctgtctt  
780  
tctattagt aggaataac actgtgtta taataacag gattcttgat acg  
840  
gaattacgt tattataaat ttatgctca gattcttgat acg  
900

(2) INFORMATION FOR SEQ ID NO:151:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 236 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) FEATURE:

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..236

(D) OTHER INFORMATION: / Ceres Seq. ID 1032070

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:151:

Met Gly Thr Thr Leu Asp Val Ser Arg Ala Glu Leu Ala Leu Val Val  
1 5 10 15  
20 Met Tyr Leu Asn Lys Ala Glu Ala Arg Asp Lys Leu Cys Arg Ala Ile  
25 30



395  
Gln Tyr Gly Ser Lys Phe Leu Ser Gly Gly Gln Pro Gly Thr Ala Gln  
35 40  
Asn Val Asp Lys Ser Thr Ser Leu Ala Arg Lys Val Phe Arg Leu Phe  
50 55  
Lys Phe Val Asn Asp Leu His Gly Leu Ile Ser Pro Val Pro Lys Gly  
65 70 75 80  
Thr Pro Leu Pro Leu Val Leu Leu Gly Lys Ser Lys Asn Ala Leu Leu  
85 90 95  
Ser Thr Phe Leu Phe Leu Asp Ctr Ile Val Trp Leu Gly Arg Ser Gly  
100 105 110  
Ile Tyr Lys Asn Lys Glu Arg Ala Glu Leu Leu Gly Arg Ile Ser Leu  
115 120 125  
Phe Cys Trp Met Gly Ser Ser Val Cys Thr Thr Leu Val Glu Val Gly  
130 135 140  
Glu Met Gly Arg Leu Ser Ser Met Lys Lys Ile Glu Lys Gly Leu  
145 150 155  
Lys Asn Gly Asn Lys Tyr Gln Asp Glu Asp Tyr Arg Ala Lys Leu Lys  
165 170 175  
Lys Ser Asn Glu Arg Ser Leu Ala Leu Ile Lys Ser Ala Met Asp Ile  
180 185 190  
Val Val Ala Ala Gly Leu Leu Gln Leu Ala Pro Thr Lys Ile Thr Pro  
195 200 205  
Arg Val Thr Gly Ala Phe Gly Phe Ile Thr Ser Ile Ile Ser Cys Tyr  
210 215 220  
Gln Leu Leu Pro Thr Arg Pro Lys Ile Lys Thr Pro  
225 230 235

(2) INFORMATION FOR SEQ ID NO:152:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 229 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..220  
(D) OTHER INFORMATION: / Ceres Seq. ID 1032071  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:152:  
Met Tyr Leu Asn Lys Ala Glu Ala Arg Asp Lys Leu Cys Arg Ala Ile  
1 5 10 15  
Gln Tyr Gly Ser Lys Phe Leu Ser Gly Gly Gln Pro Gly Thr Ala Gln  
20 25 30  
Asn Val Asp Lys Ser Thr Ser Leu Ala Arg Lys Val Phe Arg Leu Phe  
35 40 45  
Lys Phe Val Asn Asp Leu His Gly Leu Ile Ser Pro Val Pro Lys Gly  
50 55 60  
Thr Pro Leu Pro Leu Val Leu Leu Gly Lys Ser Lys Asn Ala Leu Leu  
65 70 75  
Ser Thr Phe Leu Phe Leu Asp Gln Ile Val Trp Leu Gly Arg Ser Gly  
80 85 90  
Ile Tyr Lys Asn Lys Glu Arg Ala Glu Leu Leu Gly Arg Ile Ser Leu  
95 100 105  
Phe Cys Trp Met Gly Ser Ser Val Cys Thr Thr Leu Val Glu Val Gly  
110 115 120  
Glu Met Gly Arg Leu Ser Ser Met Lys Lys Ile Glu Lys Gly Leu  
125 130 135  
Lys Asn Gly Asn Lys Tyr Gln Asp Glu Asp Tyr Arg Ala Lys Leu Lys  
140 145 150  
Lys Ser Asn Glu Arg Ser Leu Ala Leu Ile Lys Ser Ala Met Asp Ile  
155 160 165  
Val Val Ala Ala Gly Leu Leu Gln Leu Ala Pro Thr Lys Ile Thr Pro  
170 175 180 185

396  
Arg Val Thr Gly Ala Phe Gly Phe Ile Thr Ser Ile Ile Ser Cys Tyr  
195 200 205  
Gln Leu Leu Pro Thr Arg Pro Lys Ile Lys Thr Pro  
210 215 220

(2) INFORMATION FOR SEQ ID NO:153:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 769 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..769  
(D) OTHER INFORMATION: / Ceres Seq. ID 1033557  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:153:  
ggatcgctc gtgagcaga gaggtccac actatgaa gagacgga gaacaaagg  
aactgggtt ggagataaga agtgggaact tgaatggtt cctatccac ttcgttga  
gcttgggtt cgtcggttt atggaatcc gctgctcca aggcaagat ttggttct  
tctcttcg ttctctcg taagggtg agacgagtt cgtctctt cacttctc  
ctcttcgc aatgctca ttgctgtt ctccttgc cgttctct agcttctc  
tttctggt ttccattgc gtttattc aqaatcaaa ctatggact gaatggcag  
agacgag gctttgtt taacttga aaactgct tgttcaac taaggaagc  
aatcaaaa atcttgc taactatc gtttctga gaagatgag gaccatagc  
ggagagcaa cctataaagc tgaactgc agggcgct ggaacttg tcccaagtc  
aacctagca gcggaagc ggttgaag ttgttctt cttactcta tcttgaat  
tgttcatg gctttttc caccagaaga taactatga aactggttc gttgagctt  
tttctggt ttaactct tgcctctg atctgttat aatcaatg tcttggatt  
catcacat actactac tactctct attgaattg ttactact

(2) INFORMATION FOR SEQ ID NO:154:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 157 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..157  
(D) OTHER INFORMATION: / Ceres Seq. ID 1033558  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:154:  
Met Ala Ser Leu Ser Thr Ser Val Val Ala Ser Ala Ser Ser Arg Leu  
1 5 10 15  
Trp Asn Pro Ala Ala Ser Asn Gly Lys Ile Cys Val Pro Ser Ala Ser  
20 25 30  
Leu Ser Leu Arg Thr Gly Cys Arg Arg Ser Ser Ser Leu Thr Ser  
35 40 45  
Ser Ala Ser Ser Gln Leu Leu His Cys Ser Phe Leu Ser Ser Pro Val  
50 55 60  
Ser Leu Ala Ser Pro Phe Ser Gly Leu Ser Ile Ala Phe Asp Leu Ser  
65 70 75 80  
Ser Gln Thr Ser Gly Leu Asn Gly Gln Arg Arg Arg Gly Leu Val Val  
85 90 95  
Arg Ala Gly Lys Ala Ala Leu Cys Gln Thr Lys Arg Ser Arg Ser Arg  
100 105 110  
Lys Ser Leu Ala Arg Thr His Gly Phe Arg Arg Arg Met Arg Thr Thr  
115 120 125  
Ser Gly Arg Ala Thr Ile Lys Arg Arg Arg Ala Lys Gly Arg Trp Asn  
130 135 140  
Leu Cys Pro Lys Ser Asn Pro Ser Ser Gly Lys Arg Ala  
145 150 155

(2) INFORMATION FOR SEQ ID NO:155:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 157 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..157  
(D) OTHER INFORMATION: / Ceres Seq. ID 1033559  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:155:  
Met Ala Ser Leu Ser Thr Ser Val Val Ala Ser Ala Ser Ser Arg Leu  
1 5 10 15  
Trp Asn Pro Ala Ala Ser Asn Gly Lys Ile Cys Val Pro Ser Ala Ser  
20 25 30  
Leu Ser Leu Arg Thr Gly Cys Arg Arg Ser Ser Ser Leu Thr Ser  
35 40 45  
Ser Ala Ser Ser Gln Leu Leu His Cys Ser Phe Leu Ser Ser Pro Val  
50 55 60  
Ser Leu Ala Ser Pro Phe Ser Gly Leu Ser Ile Ala Phe Asp Leu Ser  
65 70 75 80  
Ser Gln Thr Ser Gly Leu Asn Gly Gln Arg Arg Arg Gly Leu Val Val  
85 90 95  
Arg Ala Gly Lys Ala Ala Leu Cys Gln Thr Lys Arg Ser Arg Ser Arg  
100 105 110  
Lys Ser Leu Ala Arg Thr His Gly Phe Arg Arg Arg Met Arg Thr Thr  
115 120 125  
Ser Gly Arg Ala Thr Ile Lys Arg Arg Arg Ala Lys Gly Arg Trp Asn  
130 135 140  
Leu Cys Pro Lys Ser Asn Pro Ser Ser Gly Lys Arg Ala  
145 150 155

397

## (1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 679 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: DNA (genomic)

(1x) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..679

(C) OTHER INFORMATION: / Ceres Seq. ID 1034688

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:155:

aaaaataaac aaaaataaac attatattac agagaataag gaaaataatga atttcaatc  
 cgaatcagga aagaataact caagctcaac accgagaggg ccagaccaca acacacccgt  
 cgaaggaacc gaagctgctc taaccaacgc gaagctcagg gaagctgcca  
 ggttgtagt gaagctgctc taaccaacgc taaccaacgc taaccaacgc  
 taagctgccc gaagctgctc ctgatatctt agagctgccc gaagataacg gaagattcga  
 tgaagaagt agacactgct agactctga caagctctgg agactctca agactctga  
 gtcgtcaac tcaacacggg ctgagtgccc tcaacacga acgagctagg ctgagctagg  
 aagtcgaac gacgagggg ctgaagaaga cgaagaagag ctgagtgagg ggttgaggg  
 ttatgcgaag atgagctcaag gttcttgaa gtgatttgat cttaattgt tgcctatcat  
 ttctgaata ataaataaa taactagat cgttgtagc tagttatgt tgcctatcat  
 atgtttagg ggaatgaga gtagtgtaa taactctgg tgatcatgaa tctatccat  
 cctgttggt attatgctc

(2) INFORMATION FOR SEQ ID NO:156:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 170 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..170

(D) OTHER INFORMATION: / Ceres Seq. ID 1034689

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:156:

Lys Lys Asn Lys Thr Lys Asn Tyr Ile Gln Glu Lys Lys Glu Lys Met  
 1 10  
 Asn Phe Ile Ser Asp Gln Val Lys Lys Ser Ser Thr Thr Pro Glu  
 20 30  
 Glu Pro Asp His Asn Lys Pro Val Glu Gly Thr Glu Thr Ala Thr Arg  
 35 40 45  
 Pro Ala Thr Asn Ala Glu Leu Met Ala Ser Ala Lys Val Val Ala Glu  
 50 55 60  
 Ala Ala Gln Ala Ala Arg Asn Glu Ser Asp Lys Lys Lys Gly  
 65 70 75  
 Lys Val Ala Gly Ala Ser Ala Asp Ile Leu Asp Ala Ala Glu Lys Tyr  
 80 85 90 95  
 Gly Lys Phe Asp Glu Lys Ser Ser Thr Gly Gln Tyr Leu Asp Lys Ala  
 100 105 110 115  
 Glu Lys Tyr Leu Asn Asp Tyr Glu Ser Ser His Ser Thr Thr Ala Gly  
 120 125  
 Gly Pro Pro Pro Thr Thr Ser Gln Ala Glu Pro Ala Ser Gln Pro Glu  
 130 135 140  
 Pro Ala Ala Lys Lys Asp Asp Glu Glu Ser Gly Gly Lys Lys Gly  
 145 150 155  
 Tyr Ala Lys Met Ala Gln Gly Phe Leu Lys  
 160 165 170

(2) INFORMATION FOR SEQ ID NO:157:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 155 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

398

## (1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 155 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..155

(C) OTHER INFORMATION: / Ceres Seq. ID 1034690

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:157:

Met Asn Phe Ile Ser Asp Gln Val Lys Lys Lys Ser Ser Thr Thr Pro  
 1 10  
 Glu Glu Pro Asp His Asn Lys Pro Val Glu Gly Thr Glu Thr Ala Thr  
 20 25 30  
 Arg Pro Ala Thr Asn Ala Glu Leu Met Ala Ser Ala Lys Val Val Ala  
 35 40 45  
 Glu Ala Ala Gln Ala Ala Arg Asn Glu Ser Asp Lys Lys Asp Lys  
 50 55 60  
 Gly Lys Val Ala Gly Ala Ser Ala Asp Ile Leu Asp Ala Ala Glu Lys  
 65 70 75 80  
 Tyr Gly Lys Phe Asp Glu Lys Ser Ser Thr Thr Gln Tyr Leu Asp Lys  
 85 90 95  
 Ala Glu Lys Tyr Leu Asn Asp Tyr Glu Ser Ser His Ser Thr Gly Ala  
 100 105 110  
 Gly Gly Pro Pro Pro Thr Ser Gln Ala Glu Pro Ala Ser Gln Pro  
 115 120 125  
 Glu Pro Ala Ala Lys Lys Asp Asp Glu Glu Ser Gly Gly Lys Lys Gly  
 130 135 140  
 Gly Tyr Ala Lys Met Ala Gln Gly Phe Leu Lys  
 145 150 155

(2) INFORMATION FOR SEQ ID NO:158:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 115 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..115

(C) OTHER INFORMATION: / Ceres Seq. ID 1034691

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:158:

Met Ala Ser Ala Lys Val Val Ala Glu Ala Ala Gln Ala Ala Arg  
 1 10  
 Asn Glu Ser Asp Lys Lys Lys Lys Val Ala Gly Ala Ser Ala  
 20 25 30  
 Asp Ile Leu Asp Ala Ala Glu Lys Tyr Gly Lys Phe Asp Glu Lys Ser  
 35 40 45  
 Ser Thr Gly Gln Tyr Leu Asp Lys Ala Glu Lys Tyr Leu Asn Asp Tyr  
 50 55 60  
 Glu Ser Ser His Ser Thr Gly Ala Gly Gly Pro Pro Pro Thr Ser  
 65 70 75 80  
 Gln Ala Glu Pro Ala Ser Gln Pro Glu Pro Ala Ala Lys Lys Asp Asp  
 85 90 95  
 Glu Glu Ser Gly Gly Lys Leu Gly Tyr Ala Lys Met Ala Gln Gly  
 100 105 110  
 Phe Leu Lys  
 115 120

(2) INFORMATION FOR SEQ ID NO:159:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 471 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: DNA (genomic)

399

## (ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..471

(C) OTHER INFORMATION: / Ceres Seq. ID 1035033

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:159:

5 atataaaa aaactatt ttgttattc tttaagttt aaacacgtc aattcatta 60  
 tatcatctat gatgaagaa cttactcaag tctgtctac tgtatgatt atcttaacca 120  
 ttcttgtct agcagtggt gcaatgagc ggcctaggaa accaaaaaa caatgtaag 180  
 acattataa gaactcaac tttgttctg cagagtgga cttatgtgt gtgaagaaga 240  
 ggggaaggg agccgtttac tgcctctct cttaaaagt ttactgtat tcatctgcc 300  
 cttacattt ttattattt caataaag agattatga taagtagta aaataatgaa 360  
 tgggtgttt atagctaac gatgatata gtgtctcta tcaataaac tatatagc 420  
 aaagtaata atataatga caatcaaac tgaagttt ttgtttatc c

(2) INFORMATION FOR SEQ ID NO:160:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 100 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..100

(D) OTHER INFORMATION: / Ceres Seq. ID 1035034

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:160:

25 Ile Lys Lys Lys Leu Ile Leu Phe Ile Leu Leu Ser Leu Lys Thr Val 15  
 1 Asn Ser Phe Ile Ser Ser Met Met Lys Lys Leu Ile Gln Val Ser Phe 20  
 Thr Val Met Ile Ile Phe Thr Ile Leu Val Leu Gly Val Val Ala Asn 25  
 Gln Gly Leu Gly Lys Pro Lys Lys Gln Cys Asn Glu Ile Leu Lys Gln 30  
 Ser Asn Cys Val Ala Ala Glu Cys Asp Ser Met Cys Val Lys Lys Arg 35  
 Gly Lys Gly Ala Gly Tyr Cys Ser Pro Ser Lys Lys Cys Tyr Cys Tyr 40  
 Tyr His Cys Pro 100

(2) INFORMATION FOR SEQ ID NO:161:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 78 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..78

(D) OTHER INFORMATION: / Ceres Seq. ID 1035035

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:161:

50 Met Met Lys Lys Leu Ile Gln Val Ser Phe Thr Val Met Ile Ile Phe 15  
 Thr Ile Leu Val Leu Gly Val Val Ala Asn Glu Gly Leu Gly Lys Pro 20  
 Lys Lys Gln Cys Asn Glu Ile Leu Lys Gln Ser Asn Cys Val Ala Ala 25  
 Glu Cys Asp Ser Met Cys Val Lys Lys Arg Gly Lys Gly Ala Gly Tyr 30  
 Cys Ser Pro Ser Lys Lys Cys Tyr Cys Tyr Tyr His Cys Pro 35  
 Tyr His Cys Pro 100

(2) INFORMATION FOR SEQ ID NO:162:

400

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 77 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..77

(D) OTHER INFORMATION: / Ceres Seq. ID 1035036

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:162:

10 Met Lys Lys Leu Ile Gln Val Ser Phe Thr Val Met Ile Ile Phe Thr 15  
 1 Ile Leu Val Leu Gly Val Val Ala Asn Glu Gly Leu Gly Lys Pro Lys 20  
 Lys Gln Cys Asn Glu Ile Leu Lys Gln Ser Asn Cys Val Ala Ala Glu 25  
 Cys Asp Ser Met Cys Val Lys Lys Arg Gly Lys Gly Ala Gly Tyr Cys 30  
 Ser Pro Ser Lys Lys Cys Tyr Cys Tyr Tyr His Cys Pro 35  
 Ser Pro Ser Lys Lys Cys Tyr Cys Tyr Tyr His Cys Pro 40  
 Ser Pro Ser Lys Lys Cys Tyr Cys Tyr Tyr His Cys Pro 45  
 Ser Pro Ser Lys Lys Cys Tyr Cys Tyr Tyr His Cys Pro 50  
 Ser Pro Ser Lys Lys Cys Tyr Cys Tyr Tyr His Cys Pro 55  
 Ser Pro Ser Lys Lys Cys Tyr Cys Tyr Tyr His Cys Pro 60  
 Ser Pro Ser Lys Lys Cys Tyr Cys Tyr Tyr His Cys Pro 65  
 Ser Pro Ser Lys Lys Cys Tyr Cys Tyr Tyr His Cys Pro 70  
 Ser Pro Ser Lys Lys Cys Tyr Cys Tyr Tyr His Cys Pro 75  
 Ser Pro Ser Lys Lys Cys Tyr Cys Tyr Tyr His Cys Pro 80  
 Ser Pro Ser Lys Lys Cys Tyr Cys Tyr Tyr His Cys Pro 85  
 Ser Pro Ser Lys Lys Cys Tyr Cys Tyr Tyr His Cys Pro 90  
 Ser Pro Ser Lys Lys Cys Tyr Cys Tyr Tyr His Cys Pro 95  
 Ser Pro Ser Lys Lys Cys Tyr Cys Tyr Tyr His Cys Pro 100

(2) INFORMATION FOR SEQ ID NO:163:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 95 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..985

(D) OTHER INFORMATION: / Ceres Seq. ID 1035071

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:163:

35 taacgaact tcttctcga atctaggtt ttgtctcta gctctctcaa gtttgcaaa 60  
 caatcagac gccacagat agcaagaaga gaaagtctgt agccagatgt gtttttacg 120  
 ctgaaccaa tgaattcta accagagagc ttgctgagga tggttactct gttttgagc 180  
 ttgcacac tctatgct actgagatca tctacagc cactctact caaacgttc 240  
 ttgtgagaa ggaagaggg attaggaat tgcactct ttacaaa agatcaaat 300  
 ttctcagga cagttgag cttatgtgt agaagttgc taacagaggt ctttgcta 360  
 ttgctcagc tggctctt ctttacaagc ttctgtgtg tctgtgtgt cgtaggtgt 420  
 gctatggt ttaagttt ttatgaga gtggagctaa agatgagaa gctatgtga 480  
 gtggaact cgtgtgca cgtgcaagt agatgaagt caagatgtt tacatgttt 540  
 cctcagtc gccaccaa gatacattg atgtcaggt taggtatgt cttctagac 600  
 aggtgtgt tggctcaag ttgaagatca tctgtactg ggaaccaa ggcacaaag 660  
 gacagatgc accatgctt atgtgtga tcatcaca accgaagaa gatcatgat 720  
 acattgacc tctcaggt ttatctcag cgtgtttgt accgaagct ccttaacca 780  
 ccaagatta tccagatg ccagttgct agagagctt ttcgtttct ggttaaac 840  
 ccaatggga tggctctt ttggagacg tactagttt tctatgtat tctaagtaa 900  
 tacttaagt cctgaggtt gaattatc aagatatga argtatgtt ttcgtatct 960  
 ttgtttaaa tttattat ctatt

(2) INFORMATION FOR SEQ ID NO:164:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 148 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..148

(D) OTHER INFORMATION: / Ceres Seq. ID 1035072

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:164:

401  
Met Ala Thr Gln Ile Ser Lys Lys Arg Lys Phe Val Ala Asp Gly Val  
1 5 10 15  
Phe Tyr Ala Glu Leu Asn Glu Val Leu Thr Arg Glu Leu Ala Glu Asp  
20 25 30  
5 Gly Tyr Ser Gly Val Glu Val Arg Val Thr Pro Met Arg Thr Glu Ile  
35 40 45  
Ile Ile Arg Ala Thr Arg Thr Gln Asn Val Leu Gly Glu Lys Gly Arg  
50 55 60  
10 Arg Ile Arg Glu Leu Thr Ser Leu Val Gln Lys Arg Phe Lys Phe Pro  
65 70 75 80  
Gln Asp Ser Val Glu Leu Tyr Ala Glu Lys Val Ala Asn Arg Gly Leu  
85 90 95  
Cys Ala Ile Ala Gln Ala Glu Ser Leu Arg Tyr Lys Leu Leu Gly G-ly  
100 105 110  
15 Leu Ala Val Arg Arg Ala Cys Tyr Gly Val Leu Arg Phe Val Met Glu  
115 120 125  
Ser Gly Ala Lys Gly Cys Glu Val Ile Val Ser Gly Lys Leu Arg Ala  
130 135 140  
Ala Arg Ala Lys  
145

(2) INFORMATION FOR SEQ ID NO:165:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 105 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..105

(D) OTHER INFORMATION: / Ceres Seq. ID 1035073

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:165:

Met Arg Thr Glu Ile Ile Arg Ala Thr Arg Thr Gln Asn Val Leu  
1 5 10 15  
Gly G-ly Lys Arg Arg Ile Arg Glu Leu Thr Ser Leu Val Gln Lys  
20 25  
Arg Phe Lys Phe Pro Gln Asp Ser Val Glu Leu Tyr Ala Glu Lys Val  
30 35 40 45  
Ala Asn Arg Gly Leu Cys Ala Ile Ala Gln Ala Glu Ser Leu Arg Tyr  
50 55 60  
Lys Leu Leu Gly Gly Leu Ala Val Arg Arg Ala Cys Tyr Gly Val Leu  
65 70 75 80  
Arg Phe Val Met Glu Ser Gly Ala Lys Gly Cys Glu Val Ile Val Ser  
85 90 95  
45 Gly Lys Leu Arg Ala Ala Arg Ala Lys  
100 105

(2) INFORMATION FOR SEQ ID NO:166:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 99 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..99

(D) OTHER INFORMATION: / Ceres Seq. ID 1035074

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:166:

Met Lys Phe Lys Asp Gly Tyr Met Val Ser Ser Gly Gln Pro Thr Lys  
1 5 10 15  
Glu Tyr Ile Asp Ala Ala Val Arg His Val Leu Leu Arg Gln Gly Val  
20 25 30

402  
Leu Gly Leu Lys Val Lys Ile Met Leu Asp Trp Asp Pro Lys Gly Lys  
35 40 45  
Gln Gly Pro Met Thr Pro Leu Pro Asp Val Val Ile Ile His Thr Pro  
50 55 60  
5 Lys Glu Asp Asp Val Tyr Ile Ala Pro Ala Gln Val Val Thr Gln Ala  
65 70 75 80  
Ala Phe Val Pro Glu Ala Pro Leu Thr Thr Thr Asp Tyr Pro Ala Met  
85 90 95  
Pro Val Ala  
10

(2) INFORMATION FOR SEQ ID NO:167:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 759 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..759

(D) OTHER INFORMATION: / Ceres Seq. ID 1376587

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:167:

a-cgcagaa accctattg cccctcttg ctaacgtca tgcctcttt gcagcgaag  
60  
cccaacaa agacgatga tctctcaga gaacatcgc aqaqatct qcaaatccct  
120  
ttccaagag ggagtagc- ttctaagaa ggatttcaat tccccaagc atccgtgat  
180  
tgatgacca aacgtcgaag tgattagct catcgaggt tcaagcca agagtagct  
240  
tagggagaca ttgctgaga tgcattata ttggttttg actaatgag gaatgagtt  
300  
cttgagaact tcttcaacc tctctcaga tgttgcctt gcctaatga gaagctctg  
360  
taagcgggt ggtctacct ttgtgccc acctggtgat cgcacaagag gccactctg  
420  
ctcgtgaga gacgtccca gatttgtga ccgtatgga taccgtgag gccacgag  
480  
tggtagag agagtgag ctcagctga cctcagcgc tcttccag aggtgaggg  
540  
tagcttgt ttggcctg gtgcagcgg ttacagtca gcacacct ctggtcagg  
600  
gtcctctga aaatttct gctatctg gcagatgga ggcagttt gttttttg  
660  
cttcaaga acaactctg tagttcagg attagtgc  
720

(2) INFORMATION FOR SEQ ID NO:168:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 202 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..202

(D) OTHER INFORMATION: / Ceres Seq. ID 1376588

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:168:

Ser Ala Glu Asn Pro Ile Ala Pro Leu Cys Ser Ser Ser Val- Pro Leu  
1 5 10 15  
Cys Ser Glu Ser Ser Thr Asn Lys Thr Met Ile Ile Ser Glu Asn Asn  
20 25 30  
Arg Arg Glu Ile Cys Lys Tyr Leu Phe Lys Glu Gly Val Cys Phe Ala  
35 40 45  
Lys Lys Asp Phe Asn Leu Pro Lys His Pro Leu Ile Asp Val Pro Asn  
50 55 60  
Leu Gln Val Ile Lys Leu Met Gln Ser Phe Lys Ser Lys Glu Tyr Val  
65 70 75 80  
Arg Glu Thr Phe Ala Trp Met His Tyr Tyr Trp Phe Leu Thr Asn Glc  
85 90 95  
Gly Ile Glu Phe Leu Arg Thr Tyr Leu Asn Leu Pro Ser Asp Val Val  
100 105 110  
Pro Ala Thr Leu Lys Ser Ala Lys Pro Gly Gly Arg Pro Phe Gly

403

115 Gly Pro Pro Gly Asp Arg Gln Arg Gly Pro Pro Arg Ser Asp Gly Asp 125  
 130 Arg Pro Arg Phe Gly Asp Arg Asp Gly Tyr Arg Gly Gly Pro Arg Gly 140  
 145 Gly Asp Gln Lys Ser Gly Ala Pro Ala Asp Phe Gln Pro Ser Phe Gln 155  
 160 Gly Gly Gly Gly Arg Pro Gly Phe Gly Arg Gly Ala Gly Gly Tyr Ser 175  
 180 Ala Ala Ala Pro Ser Gly Ser Gly Phe Pro 190  
 195 (2) INFORMATION FOR SEQ ID NO:169:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 177 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..177  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1376589

Met Ile Ile Ser Gln Asn Asn Arg Arg Gln Ile Cys Lys Tyr Leu Phe

1 5 10 15  
 Lys Gln Gly Val Cys Phe Ala Lys Lys Asp Phe Asn Leu Pro Lys His

20 Pro Leu Ile Asp Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln Ser

35 Phe Lys Ser Lys Gln Tyr Val Arg Gln Thr Phe Ala Trp Met His Tyr

50 Tyr Trp Phe Leu Thr Asn Gln Gly Ile Gln Phe Leu Arg Thr Tyr Leu

65 Asn Leu Pro Lys Asp Val Val Pro Ala Thr Leu Lys Lys Ser Ala Lys

80 Pro Gly Gly Arg Pro Phe Gly Gly Pro Pro Gly Asp Arg Gln Arg Gly

100 Pro Pro Arg Ser Asp Gly Asp Arg Pro Arg Phe Gly Asp Arg Asp Gly

115 Tyr Arg Gly Gly Pro Arg Gly Gly Asp Gln Lys Ser Gly Ala Pro Ala

130 Asp Phe Gln Pro Ser Phe Gln Gly Gly Gly Arg Pro Gly Phe Gly

145 Arg Gly Ala Gly Gly Tyr Ser Ala Ala Ala Pro Ser Gly Ser Gly Phe

160 Pro 175

(2) INFORMATION FOR SEQ ID NO:170:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 132 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..132  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1376590

Met Gln Ser Phe Lys Lys Lys Gln Tyr Val Arg Gln Thr Phe Ala Trp

1 5 10 15  
 Met His Tyr Tyr Trp Phe Leu Thr Asn Gln Gly Ile Gln Phe Leu Arg

404

20 Thr Tyr Leu Asn Leu Pro Ser Asp Val Val Pro Ala Thr Leu Lys Lys 30  
 35 Ser Ala Lys Pro Gly Tyr Arg Pro Phe Gly Gly Pro Pro Gly Asp Arg 45  
 50 Gln Arg Gly Pro Pro Arg Ser Asp Gly Asp Arg Pro Arg Phe Gly Asp 60  
 65 Arg Asp Gly Tyr Arg Gly Gly Pro Arg Gly Gly Asp Gln Lys Ser Gly 80  
 85 Ala Pro Ala Asp Phe Gln Pro Ser Phe Gln Gly Gly Gly Arg Pro 95  
 100 Gly Phe Gly Arg Gly Ala Gly Tyr Ser Ala Ala Ala Pro Ser Gly 110  
 115 Ser Gly Phe Pro 120  
 130 (2) INFORMATION FOR SEQ ID NO:171:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 645 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..645  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1378581

gatctccat ttgtgtgca ggcattgag agctctgac gcaacagca cctctcggc

ttcgaatc caataaagt gtgattgaa cgaagatcc tagattctt attggtgca

ccattatag cgggggttc agtactgca acagacttg ctggtagata tggatccaa

gcattcaag cattcaagg caagccacca agcccaaaa ttaagaaatt ttatggggc

gtttccagc ctactgac gaaaggaaa ggtctctca ttctggcgt caggggagt

gcggggcg aagattgaa ggaagcac aggaagtaa tg-agcaaa ccatccaat

gcagggtga gccatttct agctcttaag atcaatgaag ctaaaacgt gatctcagg

aaactcaaa acagggatc cgtctttga tgcacattc atcaaaaga ctctctatg

aatcactta tttatagga ctgaatgac aaatgcaga ttttatgta attataagc

ttgggtctg ttgtgacac ataaagatc gaattcttc ttctactgt tgtcgtgaa

aagaacaaa aaataaaac gcaaacaga agggcgctt adtgcg

(2) INFORMATION FOR SEQ ID NO:172:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 112 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..112  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1378582

Met Ala Thr Pro Phe Ile Ala Gly Val Ala Val Ala Ala Thr Ala Leu

1 5 10 15  
 Ala Gly Arg Tyr Gly Ile Gln Ala Trp Gln Ala Phe Lys Ala Arg Pro

20 Pro Arg Pro Lys Ile Lys Lys Phe Tyr Gln Gly Gly Phe Gln Pro Thr

35 Met Thr Lys Arg Gln Ala Ala Leu Ile Leu Gly Val Arg Gln Ser Val

50 Ala Ala Gln Lys Val Lys Gln Ala His Arg Lys Val Met Val Ala Asn

65 His Pro Asp Ala Gly Gly Ser His Phe Leu Ala Ser Lys Ile Asn Gln

80 85

405

Ala Lys Asp Val Met Leu Gly Lys Thr Lys Asn Ser Gly Ser Ala Phe  
100 105 110

(2) INFORMATION FOR SEQ ID NO:173:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 64 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..64

(D) OTHER INFORMATION: / Ceres Seq. ID 1378583

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:173:

Met Thr Lys Arg Glu Ala Ala Leu Ile Leu Gly Val Arg Glu Ser Val  
1 5 15

Ala Ala Glu Lys Val Lys Glu Ala His Arg Lys Val Met Val Ala Asn  
20 25 30

His Pro Asp Ala Gly Gly Ser His Phe Leu Ala Ser Lys Ile Asn Glu  
35 40 45

Ala Lys Asp Val Met Leu Gly Lys Thr Lys Asn Ser Gly Ser Ala Phe  
50 55 60

(2) INFORMATION FOR SEQ ID NO:174:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 803 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..803

(D) OTHER INFORMATION: / Ceres Seq. ID 1383462

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:174:

aggacacgtt tctatctaga ggaacacaa cacatcgac catggagaa tcaacagaa  
60

ccattcat accatctg ataatctca tggctagg gttggaaac tctgacctg  
120

ctcaggac agagagctt acgaacagc tcatgaact accacgtgt atccgtacg  
180

ttggaggaga cgcacaggtt ccaacaaag atgtgtgac aggtgtggc caagtataa  
240

gaagagta gaagtgtt tgcatalgg tgaagacaa agatgatct caactggta  
300

tcaagattaa cgaacacta gcgctcttc ttccttcgc ttgtcatatc acgttctta  
360

acatcaacta ctgtattctg attctgcat tactctgaa ctcaacatg gctaaagct  
420

ttgagaactt aggaayatt ygaacaaat acaactccac attctctaca caactcata  
480

agaatggac aggaggaga aaagcgac catggaag taatggatg aggagaaga  
540

gttggatgg tttgagctt ttaataatt tgtatgttc tctcatctt ttttaattac  
600

ttctcttct taatgtctt ttgttttatt attacaataa cttaaatgat atctatttat  
660

atatgaaca catgatctg ttggcatct atatgatgt aaatgaatc ctggttaatt  
720

tttaacagc aagctacat ttaattctg ttgt-gccaa ttgatctgt acgttagta  
780

atttatgga ttatgactt act

(2) INFORMATION FOR SEQ ID NO:175:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 197 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..197

(D) OTHER INFORMATION: / Ceres Seq. ID 1383463

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:175:

Asp Thr Phe Leu Ser Arg Glu Asn Thr His Arg Thr Met Glu Lys  
1 5 10

406

Ser Thr Arg Thr Leu Phe Ile Thr Ile Val Ile Thr Ser Met Leu Leu  
20 25 30

Gly Phe Gly Asn Ser Asp Leu Ala Gln Asp Arg Glu Gly Cys Thr Asn  
35 40 45

Gln Leu Ile Glu Leu Ser Thr Cys Ile Pro Tyr Val Gly Gly Asp Ala  
50 55 60

Lys Ala Pro Thr Lys Asp Cys Cys Ala Gly Phe Gly Gln Val Ile Arg  
65 70 75

Lys Ser Glu Lys Cys Val Cys Ile Leu Val Arg Asp Lys Asp Asp Pro  
85 90 95

Gln Leu Gly Ile Lys Ile Asn Ala Thr Leu Ala Ala His Leu Pro Ser  
100 105 110

Ala Cys His Ile Thr Ala Pro Asn Ile Thr Asp Cys Ile Ser Ile Leu  
115 120 125

His Leu Pro Arg Asn Ser Thr Leu Ala Lys Glu Phe Glu Asn Leu Gly  
130 135 140

Arg Ile Glu Asp Asn Tyr Asn Ser Thr Ser Pro Thr Gln Ile His Lys  
145 150 155

Asp Gly Thr Gly Gly Lys Ala Glu Pro Val Lys Ser Asn Gly Trp  
165 170 175

Lys Glu Lys Ser Trp Leu Gly Val Glu Leu Leu Ile Tyr Leu Leu Val  
180 185 190

Ser Leu Ile Phe Phe

195

(2) INFORMATION FOR SEQ ID NO:176:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 184 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: pepitid

(B) LOCATION: 1..184

(D) OTHER INFORMATION: / Ceres Seq. ID 1383464

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:176:

Met Glu Lys Ser Thr Arg Thr Leu Phe Ile Thr Ile Val Ile Thr Ser  
1 5 10

Met Leu Leu Gly Phe Gly Asn Ser Asp Leu Ala Gln Asp Arg Glu  
20 25 30

Cys Thr Asn Gln Leu Ile Glu Leu Ser Thr Cys Ile Pro Tyr Val Gly  
35 40 45

Gly Asp Ala Lys Ala Pro Thr Lys Asp Cys Cys Ala Gly Phe Gly Gln  
50 55 60

Val Ile Arg Lys Ser Glu Lys Cys Val Cys Ile Leu Val Arg Asp Lys  
65 70 75

Asp Asp Pro Gln Leu Gly Ile Lys Ile Asn Ala Thr Leu Ala Ala His  
85 90 95

Leu Pro Ser Ala Cys His Ile Thr Ala Pro Asn Ile Thr Asp Cys Ile  
100 105 110

Ser Ile Leu His Leu Pro Arg Asn Ser Thr Leu Ala Lys Glu Phe Glu  
115 120 125

Asn Leu Gly Arg Ile Glu Asp Asn Tyr Asn Ser Thr Ser Pro Thr Gln  
130 135 140

Ile His Lys Asp Gly Thr Gly Gly Lys Ala Glu Pro Val Lys Ser  
145 150 155

Asn Gly Trp Lys Glu Lys Ser Trp Leu Gly Val Glu Leu Ile Tyr  
165 170 175

Leu Leu Val Ser Leu Ile Phe Phe

180

(2) INFORMATION FOR SEQ ID NO:177:

(i) SEQUENCE CHARACTERISTICS:

407

(A) LENGTH: 168 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..168

(D) OTHER INFORMATION: / Ceres Seq. ID 1383465

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:177:

Met Leu Leu Gly Phe Gly Asn Ser Asp Leu Ala Gln Asp Arg Glu Glu

1 5 10 15

Cys Thr Asn Gln Leu Ile Glu Leu Ser Thr Cys Ile Pro Tyr Val Gly

20 25 30 35

Gly Asp Ala Lys Ala Pro Thr Lys Asp Cys Cys Ala Gly Phe Gly Gln

40 45 50 55

Val Ile Arg Lys Ser Glu Lys Cys Val Cys Ile Leu Val Arg Asp Lys

60 65 70 75

Asp Asp Pro Gln Leu Gly Ile Lys Ile Asn Ala Thr Leu Ala Ala His

80 85 90 95

Leu Pro Ser Ala Cys His Ile Thr Ala Pro Asn Ile Thr Asp Cys Ile

100 105 110 115

Ser Ile Leu His Leu Pro Arg Asn Ser Thr Leu Ala Lys Glu Phe Glu

120 125 130 135

Asn Leu Gly Arg Ile Glu Asp Asn Tyr Asn Ser Thr Ser Pro Thr Gln

140 145 150 155

Ile His Lys Asp Gly Thr Gly Gly Lys Ala Glu Pro Val Lys Ser

160 165 170 175

Asn Gly Trp Lys Glu Lys Ser Trp Leu Gly Val Glu Leu Leu Ile Tyr

180 185 190 195

Leu Leu Val Ser Leu Ile Phe

200 205 210 215

(2) INFORMATION FOR SEQ ID NO:178:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 966 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..968

(D) OTHER INFORMATION: / Ceres Seq. ID 1386215

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:178:

AAGGAGGAGG CAAAGTGGCA AAGTGGGCA TAAAGGAGG ATGCGAGG AAGTAAAG

45 50 55 60

ATTGAGGAA TCAAGTGA AATGGTGG CACTATGAT GCTATGAT GCTATGAT GCTATGAT

120 125 130 135

TGTCTTGA TCAAGTGA AATGGTGG CACTATGAT GCTATGAT GCTATGAT GCTATGAT

180 185 190 195

CATGAGAG CAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG

240 245 250 255

AATCTGTA ATGGAACG CAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG

300 305 310 315

GCTATGAT ATGGAACG CAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG

360 365 370 375

AAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG

420 425 430 435

AAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG

480 485 490 495

AAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG

540 545 550 555

AAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG

600 605 610 615

AAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG

660 665 670 675

AAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG

720 725 730 735

AAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG

780 785 790 795

AAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG GAGGAGAG

840 845 850 855

408

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 180 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..180

(D) OTHER INFORMATION: / Ceres Seq. ID 1386216

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:179:

Arg Gly Lys Gln Asn Cys Lys Asn Lys Asn Lys Lys Asn Ile Ala Gln

1 5 10 15

Glu Asn Lys Arg Phe Val Glu Ser Thr Lys Lys Met Ala Ser Thr Met

20 25 30 35

Met Thr Thr Leu Pro Gln Phe Asn Gly Leu Arg Ala Thr Lys Ile Ser

40 45 50 55

Ala Ala Pro Val Gln Gly Leu Ala Ser Val Gln Pro Met Arg Arg Lys

60 65 70 75

Gly Asn Gly Ala Leu Gly Ala Lys Cys Asp Phe Ile Gly Ser Ser Thr

80 85 90 95

Asn Leu Ile Met Val Thr Ser Thr Thr Leu Met Leu Phe Ala Gly Arg

100 105 110 115

Phe Gly Leu Ala Pro Ser Ala Asn Arg Lys Ala Thr Ala Gly Leu Arg

120 125 130 135

Leu Glu Ala Arg Asp Ser Gly Leu Gln Thr Gly Asp Pro Ala Gly Phe

140 145 150 155

Thr Leu Ala Asp Thr Leu Ala Cys Ala Ser Asn Leu Thr Pro Ile Ser

160 165 170 175

Pro Ser Ser Pro Asn Leu Glu Ser Met Lys Thr Ser Pro Met Lys Arg

180 185 190 195

Leu Ser Leu Arg Asn Arg Gln Trp Leu Leu Ser Gln Leu Glu Ile Ser

200 205 210 215

(2) INFORMATION FOR SEQ ID NO:180:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 153 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..153

(D) OTHER INFORMATION: / Ceres Seq. ID 1386217

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:180:

Met Ala Ser Thr Met Met Thr Thr Leu Pro Gln Phe Asn Gly Leu Arg

1 5 10 15

Ala Thr Lys Ile Ser Ala Ala Pro Val Gln Gly Leu Ala Ser Val Gln

20 25 30 35

Pro Met Arg Arg Lys Gly Asn Gly Ala Leu Gly Ala Lys Cys Asp Phe

40 45 50 55

Ile Gly Ser Ser Thr Asn Leu Ile Met Val Thr Ser Thr Thr Leu Met

60 65 70 75

Leu Phe Ala Gly Arg Phe Gly Leu Ala Pro Ser Ala Asn Arg Lys Ala

80 85 90 95

Thr Ala Gly Leu Arg Leu Glu Ala Arg Asp Ser Gly Leu Gln Thr Gly

100 105 110 115

Asp Pro Ala Gly Phe Thr Leu Ala Asp Thr Leu Ala Cys Ala Ser Asn

120 125 130 135

Leu Thr Pro Ile Ser Pro Ser Pro Asn Leu Glu Ser Met Lys Thr

140 145 150 155

115 120 409  
 Ser Pro Met Lys Arg Leu Ser Leu Arg Asn Arg Gln Trp Leu Leu Ser  
 130 135 140  
 Gln Leu Glu Ile Ser Arg Trp Leu Ile  
 145 150  
 (2) INFORMATION FOR SEQ ID NO:181:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 149 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..149  
 (C) OTHER INFORMATION: / Ceres Seq. ID 1386218  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:181:  
 Met Met Thr Thr Leu Pro Gln Phe Asn Gly Leu Arg Ala Thr Lys Ile  
 1 5 10 15  
 Ser Ala Ala Pro Val Gln Gly Leu Ala Ser Val Gln Pro Met Arg Arg  
 20 25 30  
 Lys Gly Asn Gly Ala Leu Gly Ala Lys Cys Asp Phe Ile Gly Ser Ser  
 35 40 45  
 Thr Asn Leu Ile Met Val Thr Ser Thr Thr Leu Met Leu Phe Ala Gly  
 50 55 60  
 Arg Phe Gly Leu Ala Pro Ser Ala Asn Arg Lys Ala Thr Ala Gly Leu  
 65 70 75 80  
 Arg Leu Glu Ala Arg Asp Ser Gly Leu Gln Thr Gly Asp Pro Ala Gly  
 85 90 95  
 Phe Thr Leu Ala Asp Thr Leu Ala Cys Ala Ser Asn Leu Thr Pro Ile  
 100 105 110  
 Ser Pro Ser Ser Pro Asn Leu Gln Ser Met Lys Thr Ser Pro Met Lys  
 115 120 125  
 Arg Leu Ser Leu Arg Asn Arg Gln Trp Leu Leu Ser Gln Leu Glu Ile  
 130 135 140  
 Ser Arg Trp Leu Ile  
 145  
 (2) INFORMATION FOR SEQ ID NO:182:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 65 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..65  
 (C) OTHER INFORMATION: / Ceres Seq. ID 2025156  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:182:  
 Met Ala Leu Val Ala Ala Gly Asn Ile Ser Val Ala Asn Met Asn Lys  
 1 5 10 15  
 Ala Val Ala Leu Leu Lys Glu Glu Ala Leu Lys Lys Ile Ile Ala Ala  
 20 25 30  
 Lys Thr Ala Lys Leu Arg Ala Lys Gly Ala Leu Thr Asn Leu Gln Asp  
 35 40 45  
 Ile Val Ile Arg Gln Ser Lys Val Thr Gly Lys Arg Lys Glu Asp Glu  
 50 55 60  
 Arg  
 65  
 (2) INFORMATION FOR SEQ ID NO:183:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 770 base pairs

41C  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..770  
 (C) OTHER INFORMATION: / Ceres Seq. ID 1388499  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:183:  
 aaaaactca ttctcaaac aattccctta actctctctc leggtgactt aagagtaaaa  
 5 10 15 20 25 30 35 40 45 50 55 60  
 aatgttagag gcagaggttg aatagatgtg ggaatcaaaa tctacggctg aagatgtttt  
 65 70 75 80 85 90 95 100 105 110 115 120  
 ctgtctctt agaagatcac aatagctctt caaagcctac cgtatgtac aggaagtga  
 125 130 135 140 145 150 155 160 165 170 175 180  
 tctctcgaa gccagctggg gcgaagtgtg aagcatactc tigtgaaat taactgtga  
 185 190 195 200 205 210 215 220 225 230 235 240  
 tggagacca aagtgtcaa aggaatcat tgaggcatt gatagaaga gaaatagat  
 245 250 255 260 265 270 275 280 285 290 295 300  
 ccagtgaggg gtgttagagg gacctctgaa ggaagagtac aatatttt cgaacacgat  
 305 310 315 320 325 330 335 340 345 350 355 360  
 gaaatgaag cctaagcaag gaggtcttgg aagtgttg aagtgaact tgaacatga  
 365 370 375 380 385 390 395 400 405 410 415 420  
 gagaattgat gaaaagttgg ctacactaga gagctcctt cagttctcg tgaatgtgt  
 425 430 435 440 445 450 455 460 465 470 475 480  
 taagagatt gaccaatacc tatgtctga aggttaaggg accttctcc ctggtctatg  
 485 490 495 500 505 510 515 520 525 530 535 540  
 tctggggag ttattgtatg cgtacgtga agactcttc ttgtttta tatgtcttc  
 545 550 555 560 565 570 575 580 585 590 595 600  
 tcaataatg gatacttga gattagaag agaggtatcc aagtgtcta taactcatg  
 605 610 615 620 625 630 635 640 645 650 655 660  
 cttaaatgg gaatttcat- aacttttaa gtgtcttatg aatttccat atatgtatg  
 665 670 675 680 685 690 695 700 705 710 715 720  
 ttttttgg ttgaataac atattaat gcaacataac catgttttt  
 (2) INFORMATION FOR SEQ ID NO:184:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 151 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..151  
 (C) OTHER INFORMATION: / Ceres Seq. ID 1388500  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:184:  
 Met Val Gln Ala Glu Val Glu Val Asp Val Glu Ile Lys Ser Thr Ala  
 1 5 10 15  
 Asp Lys Phe Phe Met Phe Ser Arg Arg Ser Gln His Ala Ser Lys Ala  
 20 25 30  
 Thr Arg Tyr Val Gln Gly Cys Asp Leu Leu Glu Gly Glu Trp Gly Glu  
 35 40 45  
 Val Gly Ser Ile Leu Leu Trp Lys Leu Thr Val Asp Gly Glu Pro Lys  
 50 55 60  
 Val Ser Lys Asp Met Ile Glu Ala Ile Asp Met Lys Met Asn Met Ile  
 65 70 75 80 85 90 95 100 105 110 115 120  
 Gln Trp Arg Val Leu Glu Gly Pro Leu Lys Glu Glu Tyr Asn Ile Phe  
 125 130 135 140 145 150  
 Ser Lys Thr Met Lys Val Ser Pro Lys Gln Gly Ser Gly Ser Val  
 155 160 165 170 175 180 185 190 195 200 205 210  
 Val Lys Trp Asn Leu Lys Tyr Glu Arg Ile Asp Glu Lys Val Ala His  
 215 220 225 230 235 240 245 250 255 260 265 270  
 Leu Glu Arg Leu Leu Gln Phe Val Glu Cys Val Asn Glu Ile Asp  
 275 280 285 290 295 300 305 310 315 320 325 330  
 Gln Tyr Leu Leu Ser Glu Gly  
 335 340 345 350 355 360 365 370 375 380 385 390  
 (2) INFORMATION FOR SEQ ID NO:185:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 131 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:



411

(A) NAME/KEY: peptide  
(B) LOCATION: 1..131  
(D) OTHER INFORMATION: / Ceres Seq. ID 1388501  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:185:  
5 Met Phe Ser Arg Arg Ser Gln His Ala Ser Lys Ala Thr Arg Tyr Val  
1 5 10 15  
Gln Gly Cys Asp Leu Leu Glu Gly Glu Trp Gly Glu Val Gly Ser Ile  
20 25 30  
Leu Leu Trp Lys Leu Thr Val Asp Gly Glu Pro Lys Val Ser Lys Asp  
35 40 45  
Met Ile Glu Ala Ile Asp Met Lys Met Asn Met Ile Gln Trp Arg Val  
50 55 60  
Leu Glu Gly Pro Lys Glu Glu Tyr Asn Ile Phe Ser Lys Thr Met  
65 70 75 80  
Lys Val Ser Pro Lys Gln Gly Ser Gly Ser Val Val Lys Trp Asn  
85 90 95  
Leu Lys Tyr Glu Arg Ile Asp Glu Lys Val Ala His Leu Glu Arg Leu  
100 105 110  
Leu Gln Phe Val Glu Cys Val Asn Glu Ile Asp Gln Tyr Leu Leu  
115 120 125  
Ser Glu Gly  
130

(2) INFORMATION FOR SEQ ID NO:186:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 83 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..83

(D) OTHER INFORMATION: / Ceres Seq. ID 1388502

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:186:

35 Met Ile Glu Ala Ile Asp Met Lys Met Asn Met Ile Gln Trp Arg Val  
1 5 10 15  
Leu Glu Gly Pro Leu Lys Glu Glu Tyr Asn Ile Phe Ser Lys Thr Met  
20 25 30  
Lys Val Ser Pro Lys Gln Gly Gly Ser Gly Ser Val Val Lys Trp Asn  
35 40 45  
Leu Lys Tyr Glu Arg Ile Asp Glu Lys Val Ala His Leu Glu Arg Leu  
50 55 60  
Leu Gln Phe Phe Val Glu Cys Val Asn Glu Ile Asp Gln Tyr Leu Leu  
65 70 75  
Ser Glu Gly  
80

(2) INFORMATION FOR SEQ ID NO:187:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 752 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..752

(D) OTHER INFORMATION: / Ceres Seq. ID 1388519

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:187:

60 aatctacgc gctttaccca tcttctctct agattctctc ctccactatt cgaataatg  
120 gcgaactgg tgaattgtt gttctctgt atctaccgg cgaatgcgt ggcgcgaagg  
180 aggggaata ttgaaagaa tacaattggg gtccaagga tacaattctg tgaatttgc  
240 aaattcggt tgaactccc tgaattctct tacttctatc cgggtgcac ggtgaagata

412

tcatgcaag acaggaagac aatggaagag gttacacag acaagctgt atcagacaaa  
gaagaaagt ataagttcat tgtccacgac gatcacagag accagatgtg cgaattttg  
ctgtgaaaa gctcgataa aacctgtctt aaactctcgg ttgacgtga gaactctgt  
420 gtgatttga accattacag tggatgtcgc tgcagatca gacatgataa caacatggga  
480 ttgagaag aagtgtgtga tgtgtctgc tgcgttctgt ttccagaagta taaggttgat  
540 gaagtaggg atgatatata aaacctatc taattctct tttaactctt atgattctgt  
600 ttctcttca ttaatgagtt tccagtatg gaagagatat attgtattt ttgtgtaac  
660 ttttgtgt tctttagaga tttgtactct ggtgatcga taactatcgg ttgtgttaag  
720 cttcttatc attgattgt cattctctg cc

(2) INFORMATION FOR SEQ ID NO:188:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 190 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..190

(D) OTHER INFORMATION: / Ceres Seq. ID 1388520

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:188:

20 Asn Ser Thr Arg Phe His His Leu Leu Pro Arg Phe Ala Pro Pro Leu  
1 5 10 15  
Phe Glu Lys Met Ala Lys Leu Val Met Leu Leu Val Leu Cys Ile Leu  
20 25 30  
Pro Ala Ile Ala Met Ala Ala Arg Arg Gly Asn Ile Gly Lys Asn Thr  
35 40 45  
Met Val Val Gln Gly Ser Thr Tyr Cys Asp Ile Cys Lys Phe Gly Phe  
50 55 60  
Glu Thr Pro Glu Ser Ser Tyr Phe Ile Pro Gly Ala Thr Val Lys Leu  
65 70 75 80  
Ser Cys Lys Asp Arg Lys Thr Met Glu Glu Val Tyr Thr Asp Lys Ala  
85 90 95  
Val Ser Asp Lys Glu Gly Lys Tyr Lys Phe Ile Val His Asp Asp His  
100 105 110  
Arg Asp Gln Met Cys Asp Val Leu Leu Val Lys Ser Ser Asp Lys Thr  
115 120 125  
Cys Ser Lys Ile Ser Val Gly Arg Glu Lys Ser Arg Val Ile Leu Asn  
130 135 140  
His Tyr Ser Gly Ile Ala Ser Gln Ile Arg His Ala Asn Asn Met Gly  
145 150 155 160  
Phe Glu Lys Glu Val Ser Asp Val Phe Cys Ser Ala Leu Phe Gln Lys  
165 170 175  
Tyr Met Val Asp Glu Asp Glu Asp Asp Ile Lys Asn His Leu  
180 185 190

(2) INFORMATION FOR SEQ ID NO:189:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 171 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..171

(D) OTHER INFORMATION: / Ceres Seq. ID 1388521

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:189:

Met Ala Lys Leu Val Met Leu Leu Val Leu Cys Ile Leu Pro Ala Ile  
1 5 10 15  
Ala Met Ala Ala Arg Arg Gly Asn Ile Gly Lys Asn Thr Met Val Val  
20 25 30  
Gln Gly Ser Thr Tyr Cys Asp Ile Cys Lys Phe Gly Phe Glu Thr Pro

413

35 Glu Ser Ser Tyr Phe Ile Pro Gly Ala Thr Val Lys Leu Ser Cys Lys  
50 55 60  
5 Asp Arg Lys Thr Met Glu Glu Val Tyr Thr Asp Lys Ala Val Ser Asp  
65 70 75  
Lys Glu Gly Lys Tyr Lys Phe Ile Val His Asp Asp His Arg Asp Glu  
85 90 95  
Met Cys Asp Val Leu Leu Val Lys Ser Ser Asp Lys Thr Cys Ser Lys  
100 105 110  
10 Ile Ser Val Gly Arg Glu Lys Ser Arg Val Ile Leu Asn His Tyr Ser  
115 120 125  
Gly Ile Ala Ser Gln Ile Arg His Ala Asn Asn Met Gly Phe Glu Lys  
130 135 140  
Glc Val Ser Asp Val Phe Cys Ser Ala Leu Phe Gln Lys Tyr Met Val  
145 150 155  
15 Asp Glu Asp Glu Asp Asp Ile Lys Asn His Leu  
165 170

(2) INFORMATION FOR SEQ ID NO:190:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 166 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..166

(D) OTHER INFORMATION: / Ceres Seq. ID 1388522

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:190:

30 Met Leu Leu Val Leu Cys Ile Leu Pro Ala Ile Ala Met Ala Arg  
1 5 10 15  
Arg Gly Asn Ile Gly Lys Asn Thr Met Val Val Gln Gly Ser Thr Tyr  
20 25 30  
35 Cys Asp Ile Cys Lys Phe Gly Phe Glu Thr Pro Glu Ser Ser Tyr Phe  
35 40 45  
Ile Pro Gly Ala Thr Val Lys Leu Ser Cys Lys Asp Arg Lys Thr Met  
50 55 60  
Glu Glu Val Tyr Thr Asp Lys Ala Val Ser Asp Lys Glu Gly Lys Tyr  
65 70 75  
40 Lys Phe Ile Val His Asp Asp His Arg Asp Gln Met Cys Asp Val Leu  
85 90 95  
Leu Val Lys Ser Ser Asp Lys Thr Cys Ser Lys Ile Ser Val Gly Arg  
100 105 110  
Glu Lys Ser Arg Val Ile Leu Asn His Tyr Ser Gly Ile Ala Ser Glc  
115 120 125  
45 Ile Arg His Ala Asn Asn Met Gly Phe Glu Lys Glu Val Ser Asp Val  
130 135 140  
Phe Cys Ser Ala Leu Phe Gln Lys Tyr Met Val Asp Glu Asp Glu Asp  
145 150 155  
50 Asp Ile Lys Asn His Leu  
165 170

(2) INFORMATION FOR SEQ ID NO:191:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 658 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..658

(D) OTHER INFORMATION: / Ceres Seq. ID 1388563

414

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:191:  
aagaagtttt cttcatatcc aaaaagaaa aagaagtttcc caaaaacaaa aaacttaccc 60  
aagtttagat atgatatctg aattatata aagattgtcc aaattcaatt tcccaacacc 120  
cacaagaatc cgaaccacaaa cttcaaaaga caaagaagga cgaaggtgat gagacgaag 180  
atgacctcag ctgcagcaga cccatccccc aagaacacaa gattccgcc gttatagact 240  
ctccactccc tccgcagga aacacccggc taccacgcgc agcacgcgc gttacgagcg 300  
ctctatgat cagatgptgc aagagaagc tttatagtc gattgtgag atatcatga 360  
atcgggaaga gatgaacctc ttctctctc cgtctacaa tgagacgicg actacgacta 420  
aaccgggagc aagttacctc tattgtctc gaagatgag ctaattcaa tattttacali 480  
tttttaagt ttacttgaa atattgtgaa attaattatc tgttggtgt cgttttcaaa 540  
cattttaat ttacttgaa atattgtgag ataatatttc gcaacgcgc atattaatt 600  
cgatggagcg ggtcgactgt gtaattgag taataatga aatttaattc caattgcg

(2) INFORMATION FOR SEQ ID NO:192:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 55 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..55

(D) OTHER INFORMATION: / Ceres Seq. ID 1388564

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:192:

25 Arg Ser Phe Leu Tyr Ile His Lys Arg Lys Lys Val Ser Gln Lys Thr  
1 5 10 15  
Lys Thr Tyr Thr Ser Leu Asp Met Asp Leu Glu Leu Leu Gln Asp Leu  
20 25 30  
30 Ser Lys Phe Asn Phe Pro Thr Pro Ile Lys Ile Arg Ser Lys Thr Ser  
35 40 45  
Lys Asp Lys Glu Gly Arg Arg  
50 55

(2) INFORMATION FOR SEQ ID NO:193:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 50 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..50

(D) OTHER INFORMATION: / Ceres Seq. ID 1388565

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:193:

45 Met Ile Arg Ser Cys Lys Arg Lys Leu Leu Val Ser Thr Cys Glu Ile  
1 5 10 15  
Ile Met Asn Arg Glu Glu Ile Asp Arg Phe Phe Ser Ser Val Tyr Asn  
20 25 30  
50 Glc Thr Ser Thr Thr Ala Lys Arg Arg Arg Ser Tyr Pro Tyr Cys Ser  
35 40 45  
Arg Arg  
50 55

(2) INFORMATION FOR SEQ ID NO:194:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..39

## 415

(D) OTHER INFORMATION: / Ceres Seq. ID 1388566

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:194:

Met Arg Leu Asn Ser Ile Phe Thr Phe Thr Gly Asn  
 1 5 10 15  
 Ile Val Lys Leu Ile Ile Cys Trp Cys Ser Val Leu Asn Ile Phe Asn  
 20 25 30  
 Leu Val Met Asn Met Asp Gly  
 35

(2) INFORMATION FOR SEQ ID NO:195:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 866 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..866

(D) OTHER INFORMATION: / Ceres Seq. ID 1388793

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:195:

aacgtgtgg gctaatctca cgcgtttcca tcactctct ctagattccg atctccaat  
 60  
 gtttaaat tcaatattt cttcttcac tctctctct ctaacttcaa atccataatc  
 120  
 gctcaactc tctctcttc tagattcgaa aaatggcga aactgtgat gtgtgtgtt  
 180  
 ctttgtatt taccggcgtt agcatggcg gcaaggagg gaaatattgg caagaataa  
 240  
 atgggttcc aaggttagac ctactgtac attgcaat tgggtttcga gactctgaa  
 300  
 tctctact taccctcgg tgaacggtt agctatcat gcaagacag gaagacaat  
 360  
 gaaaggttt acacagcaa agctgatca gacaagaag gaaagtataa gtctatg-c  
 420  
 caagacata acagagacca gatggcat gtttgtctg tgaagaatc gataaacc  
 480  
 tgcctaaaa tctcgtgag acgtgagat tctgtgtga cttgaacca ttaacgtgc  
 540  
 atgctctc agatcacga tgcataaac atgggtttg agaagaagt ggtgtgtg  
 600  
 tctctgtt cttgttcca gaggatgat ttgtgatga atgagatga tattaatac  
 660  
 caactaat cctctgtt aactataga tctgtgtt tcttcataa tgatttcga  
 720  
 gtaaggag acatatatt gtttgttt gattactat ttttgtct tagagattt  
 780  
 gactctgtt acgagaaac tactgttt tgaagctt ttataattg atatgtatt  
 840  
 tcttgttta ggaagtgtt tttatt

(2) INFORMATION FOR SEQ ID NO:196:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 171 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..171

(D) OTHER INFORMATION: / Ceres Seq. ID 1388794

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:196:

Met Ala Lys Leu Val Met Leu Leu Val Leu Cys Ile Leu Pro Ala Ile  
 1 5 10 15  
 Ala Met Ala Ala Arg Arg Gly Asn Ile Gly Lys Asn Thr Met Val Val  
 20 25 30  
 Gln Gly Ser Thr Tyr Cys Asp Ile Cys Lys Phe Gly Phe Glu Thr Pro  
 35 40 45  
 Glu Ser Ser Tyr Phe Ile Pro Gly Ala Thr Val Lys Leu Ser Cys Lys  
 50 55 60  
 Asp Arg Lys Thr Met Glu Glu Val Tyr Thr Asp Lys Ala Val Ser Asp  
 65 70 75  
 Lys Glu Gly Lys Tyr Lys Phe Ile Val His Asp Asp His Arg Asp Gln  
 80 85 90 95  
 Met Cys Asp Val Leu Leu Val Lys Ser Ser Asp Lys Thr Cys Ser Lys  
 100 105 110  
 Ile Ser Val Gly Arg Glu Lys Ser Arg Val Ile Leu Asn His Tyr Ser

## 416

Gly Ile Ala Ser Gln Ile Arg His Ala Asn Asn Met Gly Phe Glu Lys  
 115 120 125  
 Glu Val Ser Asp Val Phe Cys Ser Ala Leu Phe Gln Lys Tyr Met Val  
 130 135 140  
 5 145 150 155 160  
 Asp Glu Asp Glu Asp Ile Lys Asn His Leu  
 165 170

(2) INFORMATION FOR SEQ ID NO:197:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 166 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..166

(D) OTHER INFORMATION: / Ceres Seq. ID 1388795

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:197:

Met Leu Leu Val Leu Cys Ile Leu Pro Ala Ile Ala Met Ala Ala Arg  
 1 5 10 15  
 Arg Gly Asn Ile Gly Lys Asn Thr Met Val Val Gln Gly Ser Thr Tyr  
 20 25 30  
 Cys Asp Ile Cys Lys Phe Gly Phe Glu Thr Pro Cys Ser Ser Tyr Phe  
 35 40 45  
 Ile Pro Gly Ala Thr Val Lys Leu Ser Cys Lys Asp Arg Lys Thr Met  
 50 55 60  
 Glu Glu Val Tyr Thr Asp Lys Ala Val Ser Asp Lys Glu Gly Lys Tyr  
 65 70 75 80  
 Lys Phe Ile Val His Asp His Arg Asp Gln Met Cys Asp Val Leu  
 85 90 95  
 Leu Val Lys Ser Ser Asp Lys Thr Cys Ser Lys Ile Ser Val Gly Arg  
 100 105 110  
 Glu Lys Ser Arg Val Ile Leu Asn His Tyr Ser Gly Ile Ala Ser Gln  
 115 120 125  
 Ile Arg His Ala Asn Asn Met Gly Phe Glu Lys Glu Val Ser Asp Val  
 130 135 140  
 Phe Cys Ser Ala Leu Phe Gln Lys Tyr Met Val Asp Glu Asp Glu Asp  
 145 150 155 160  
 Asp Ile Lys Asn His Leu  
 165

(2) INFORMATION FOR SEQ ID NO:198:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 154 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..154

(D) OTHER INFORMATION: / Ceres Seq. ID 1388796

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:198:

Met Ala Ala Arg Arg Gly Asn Ile Gly Lys Asn Thr Met Val Val Gln  
 1 5 10 15  
 Gly Ser Thr Tyr Cys Asp Ile Cys Lys Phe Gly Phe Glu Thr Pro Glu  
 20 25 30  
 Ser Ser Tyr Phe Ile Pro Gly Ala Thr Val Lys Leu Ser Cys Lys Asp  
 35 40 45  
 Arg Lys Thr Met Glu Glu Val Tyr Thr Asp Lys Ala Val Ser Asp Lys  
 50 55 60  
 Glu Gly Lys Tyr Lys Phe Ile Val His Asp Asp His Arg Asp Gln Met

417 75 80

65 Cys Asp Val Leu Val Lys Ser Ser Asp Lys Thr Cys Ser Lys Ile 85 95

5 Ser Val Gly Arg Glu Lys Ser Arg Val Ile Leu Asn His Tyr Ser Gly 100 110

115 Ala Ser Gln Ile Arg His Ala Asn Met Gly Phe Glu Lys Glu 120 125

Val Ser Asp Val Phe Cys Ser Ala Leu Phe Gln Lys Tyr Met Val Asp 130 140

10 Glu Asp Glu Asp Asp Ile Lys Asn His Leu 135 150

145 (2) INFORMATION FOR SEQ ID NO:199:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 533 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(1x) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..533

(D) OTHER INFORMATION: / Ceres Seq. ID :392041

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:199:

aaccataca caatacace aaaaaaaa aaacgaag taacatqaga tctctctat

120 taccgtgtg cttgttctt gctttacct ggggtgaagc agcgtgtct tgcacacgg

25 tgaatcgga tctttacct tgtttacct acgtgactca gggcgacgg gcccccacc

180 tctgtcgca cgtctcaca acctcaca gctgactca acctctatg gacgtcagg

240 gggctcgtg ttgatcaaa tctgctatt gaggactcac tctctctct agaacctcc

300 aaatgctt ggaattcgt tclaaatgt ggtgtgact cccttacaag ttcagccct

360 ccaactgag cgcagatc cagtgaca agcagaaa cttaaaggaa gctactaca

420 gaactaat acccaataa ttaataatg aggcattgg ttgcactgt gtaattgat

480 cagtagta ttgtattt gaatttcta atatcagag gcaattatct ctg

(2) INFORMATION FOR SEQ ID NO:200:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 113 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..113

(D) OTHER INFORMATION: / Ceres Seq. ID 1392042

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:200:

Met Arg Ser Leu Leu Ala Val Cys Leu Val Leu Ala Leu His Cys

1 5 10 15

Gly Glu Ala Ala Val Ser Cys Asn Thr Val Ile Ala Asp Leu Tyr Pro

20 25 30

Cys Leu Ser Tyr Val Thr Gln Gly Gly Pro Val Pro Thr Leu Cys Cys

35 40 45

Asn Gly Leu Thr Thr Leu Lys Ser Gln Ala Gln Thr Ser Val Asp Arg

50 55

Gln Gly Val Cys Arg Cys Ile Lys Ser Ala Ile Gly Gly Leu Thr Leu

65 70 75

Ser Pro Arg Thr Ile Gln Asn Ala Leu Glu Leu Pro Ser Lys Cys Gly

80 85 90

Val Asp Leu Pro Tyr Lys Phe Ser Pro Ser Thr Asp Cys Asp Ser Ile

100 105 110

Gln

(2) INFORMATION FOR SEQ ID NO:201:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1028 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(1x) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..1028

(D) OTHER INFORMATION: / Ceres Seq. ID 1393342

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:201:

aatattct taattgct tatcgagaa ggcattact gttttcagg cgaactcgt

60 ttaattctt cgaactcgt cgaactcgt cgaactcgt cgaactcgt cgaactcgt

120 agtaccttg agaccacag agctgagctt ggtgtgttag ttgtgtatt gaataagcg

180 gagcgagag acaagatttg tgaagctatt caataggtt ccaatgctt agttgtgag

240 caactcgca ctgtcaaaa tctgacaaag aacacacgt tgcctagaa agtttcgt

300 ccttcaat ttgaatga tcttctgct cctcttctc cttctccaa agggactcca

360 cctcgttg ttctgtcg aaactcaaa aacgcgtgc tgcacaggt cttgtctca

420 gatcaaat ttgtgttg caggaatggg attacaagg acaagaacg tgcgtgatt

480 ctggacgta taccattt ctgttgatg ggttctcgg ttgcacatc ctgtgtgag

540 gttgggagc ttgttagct gtacacgca attaagaag tagaanaag atagggaac

600 aaggataac accagaatga gcaatcgtt gcaaaagtag agaatcaaa cgaagctca

660 ttgctctga tcaagcagg ggtgattga gttgtgctt tggattgct tcaattgct

720 ccaagcag tcaatcccg agtcacagg gcttctgct tgcctcctc actatctct

780 tgaatcag tatgcctc acatcccaag tccaaatgg tgcataaaa ggaacacag

840 gaaagcgaa ctatctctc tgcataatg ttccatcat agttgtgctc tgtttacta

900 ttaattggt ttaattgt ctgtctgac tacttaatg ttgtgttag attctatgt

960 atataagata atgataaaca ttgcagcttc tctagaataa taagaacaga taagcattac

1020 gaggctgc

418

(A) LENGTH: 1028 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(1x) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..1028

(D) OTHER INFORMATION: / Ceres Seq. ID 1393342

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:201:

aatattct taattgct tatcgagaa ggcattact gttttcagg cgaactcgt

60 ttaattctt cgaactcgt cgaactcgt cgaactcgt cgaactcgt cgaactcgt

120 agtaccttg agaccacag agctgagctt ggtgtgttag ttgtgtatt gaataagcg

180 gagcgagag acaagatttg tgaagctatt caataggtt ccaatgctt agttgtgag

240 caactcgca ctgtcaaaa tctgacaaag aacacacgt tgcctagaa agtttcgt

300 ccttcaat ttgaatga tcttctgct cctcttctc cttctccaa agggactcca

360 cctcgttg ttctgtcg aaactcaaa aacgcgtgc tgcacaggt cttgtctca

420 gatcaaat ttgtgttg caggaatggg attacaagg acaagaacg tgcgtgatt

480 ctggacgta taccattt ctgttgatg ggttctcgg ttgcacatc ctgtgtgag

540 gttgggagc ttgttagct gtacacgca attaagaag tagaanaag atagggaac

600 aaggataac accagaatga gcaatcgtt gcaaaagtag agaatcaaa cgaagctca

660 ttgctctga tcaagcagg ggtgattga gttgtgctt tggattgct tcaattgct

720 ccaagcag tcaatcccg agtcacagg gcttctgct tgcctcctc actatctct

780 tgaatcag tatgcctc acatcccaag tccaaatgg tgcataaaa ggaacacag

840 gaaagcgaa ctatctctc tgcataatg ttccatcat agttgtgctc tgtttacta

900 ttaattggt ttaattgt ctgtctgac tacttaatg ttgtgttag attctatgt

960 atataagata atgataaaca ttgcagcttc tctagaataa taagaacaga taagcattac

1020 gaggctgc

(2) INFORMATION FOR SEQ ID NO:202:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 235 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..235

(D) OTHER INFORMATION: / Ceres Seq. ID 1393343

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:202:

Met Ser Thr Leu Glu Thr Thr Arg Ala Glu Leu Gly Leu Val Val Val

1 5 10 15

Tyr Leu Asn Lys Ala Glu Ala Arg Asp Lys Ile Cys Arg Ala Ile Gln

20 25 30

Tyr Gly Ser Lys Phe Leu Ser Asp Gly Gln Pro Gly Thr Ala Gln Asn

35 40 45

Val Asp Lys Asn Thr Ser Leu Ala Arg Lys Val Phe Arg Leu Phe Lys

50 55 60

Phe Val Asn Asp Leu His Ala Leu Ile Ser Pro Val Pro Lys Gly Thr

65 70 75

Pro Leu Pro Leu Val Leu Leu Gly Lys Ser Lys Asn Ala Leu Leu Ser

80 85 90

Thr Phe Leu Phe Leu Asp Gln Ile Val Trp Leu Gly Arg Thr Gly Ile

95 100 105

Tyr Lys Asp Lys Glu Arg Ala Glu Ile Leu Gly Arg Ile Ser Leu Phe

110 115 120

Cys Trp Met Gly Ser Ser Val Cys Thr Ser Leu Val Glu Val Gly Glu

125 130 135

Leu Gly Arg Leu Ser Ile Lys Lys Leu Glu Lys Glu Ile Gly

140 145 150

Asn Lys Asp Lys His Gln Asn Glu Glu Tyr Arg Ala Lys Val Glu Lys

155 160 165

170

419

Ser Asn Glu Arg Ser Leu Ala Leu Ile Lys Ala Gly Met Asp Val Val  
180 185 190  
Val Ala Phe Gly Leu Leu Gln Leu Ala Pro Lys Lys Val Thr Pro Arg  
195 200 205  
Val Thr Gly Ala Phe Gly Phe Ala Ser Ser Leu Ile Ser Cys Tyr Gln  
210 215 220  
Leu Leu Pro Ser His Pro Lys Ser Lys Met Val  
225 230 235

(2) INFORMATION FOR SEQ ID NO:203:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 653 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..653

(D) OTHER INFORMATION: / Ceres Seq. ID 1393556

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:203:

tcttccaat ccaacagct tcatgttt attctcaat caaagcagct caaatcaagc 60  
tcgaagat gatcacatg acctctact gggaatcaa agccacaatt cctctgatt 120  
tcgaaac tgactcatg ctaagtaca tctcacttt aatcgctgc ttctctctt 180  
cgctttcta tcaactctc gagaatgcc gctccaatt caatccctt tctctctcc 240  
cggtgtccc tcaacgctt cgtctctct cggcgcttc cggcctctt atctctaat 300  
ccgttacag atccgcgct aagctgctt cgtttcttt ttccagctc aacgcagcga 360  
tcggtactt gctgactgt gaagctatg cttcaacgg aggtgtttc atcgcat-g 420  
tcgtcgatt aacgcgcga taccgtgtt tttagatctga tgacgcggt gctataccg 480  
caacgatga tctgtctga tgcgttatg aatgatgat aataaacta aaatggaaa 540  
aatccaaa aggtgtg- tgataatga tttgttat ttgtttgt tctgtatgac 600  
tattgtttt gaacaaat atccaactt aaatgtttt cttgatatga tgc

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 168 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..168

(D) OTHER INFORMATION: / Ceres Seq. ID 1393557

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:204:

Phe Pro Ile Pro Thr Ala Ser Ser Phe Thr Leu Pro Ile Lys Ala Ala  
1 5 10 15  
Gln Ile Lys Leu Ala Lys Met Met His Met Thr Phe Tyr Thr Gly Ile  
20 25 30  
Lys Ala Thr Ile Leu Phe Asp Phe Thr Lys Thr Asp Ser Thr Leu Ser  
35 40 45  
Tyr Ile Leu Thr Leu Ile Ala Cys Phe Val Phe Ser Ala Phe Tyr Gln  
50 55 60  
Tyr Leu Glu Asn Arg Arg Ile Gln Phe Lys Ser Leu Ser Ser Ser Arg  
65 70 75 80  
Arg Ala Pro Pro Pro Arg Ser Ser Ser Gly Val Ser Ala Pro Leu  
85 90 95  
Ile Pro Lys Ser Gly Thr Arg Ser Ala Ala Lys Ala Ser Val Leu  
100 105 110  
Leu Phe Gly Val Asn Ala Ala Ile Gly Tyr Leu Leu Met Leu Ala Ala  
115 120 125  
Met Ser Phe Asn Gly Gly Val Phe Ile Ala Ile Val Val Gly Leu Thr  
130 135 140  
Ala Gly Tyr Ala Val Phe Arg Ser Asp Gly Gly Ala Asp Thr Ala

420

Thr Asp Asp Pro Cys Pro Cys Ala  
150 155 160

(2) INFORMATION FOR SEQ ID NO:205:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 146 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..146

(D) OTHER INFORMATION: / Ceres Seq. ID 1393558

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:205:

Met Met His Met Thr Phe Tyr Trp Gly Ile Lys Ala Thr Ile Leu Phe  
1 5 10 15  
Asp Phe Thr Lys Thr Asp Ser Trp Leu Ser Tyr Ile Leu Thr Leu Ile  
20 25 30  
Ala Cys Phe Val Phe Ser Ala Phe Tyr Gln Tyr Leu Glu Asn Arg Arg  
35 40 45  
Ile Gln Phe Lys Ser Leu Ser Ser Arg Arg Ala Pro Pro Pro Pro  
50 55 60  
Arg Ser Ser Ser Gly Val Ser Ala Pro Leu Ile Pro Lys Ser Gly Thr  
65 70 75  
Arg Ser Ala Ala Lys Ala Ala Ser Val Leu Leu Phe Gly Val Asn Ala  
80 85 90  
Ala Ile Gly Tyr Leu Leu Met Leu Ala Ala Met Ser Phe Asn Gly Gly  
95 100 105  
Val Phe Ile Ala Ile Val Val Gly Leu Thr Ala Gly Tyr Ala Val Phe  
110 115 120  
Arg Ser Asp Asp Gly Gly Ala Asp Thr Ala Thr Asp Asp Pro Cys Pro  
125 130 135 140  
Cys Ala

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 145 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..145

(D) OTHER INFORMATION: / Ceres Seq. ID 1393559

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:206:

Met His Met Thr Phe Tyr Trp Gly Ile Lys Ala Thr Ile Leu Phe Asp  
1 5 10 15  
Phe Thr Lys Thr Asp Ser Trp Leu Ser Tyr Ile Leu Thr Leu Ile Ala  
20 25 30  
Cys Phe Val Phe Ser Ala Phe Tyr Gln Tyr Leu Glu Asn Arg Arg Ile  
35 40 45  
Gln Phe Lys Ser Leu Ser Ser Arg Arg Ala Pro Pro Pro Pro Arg  
50 55 60  
Ser Ser Ser Gly Val Ser Ala Pro Leu Ile Pro Lys Ser Gly Thr Arg  
65 70 75  
Ser Ala Ala Lys Ala Ala Ser Val Leu Leu Phe Gly Val Asn Ala Ala  
80 85 90 95  
Ile Gly Tyr Leu Leu Met Leu Ala Ala Met Ser Phe Asn Gly Gly Val  
100 105 110  
Phe Ile Ala Ile Val Val Gly Leu Thr Ala Gly Tyr Ala Val Phe Arg

421

115 120 125  
Ser Asp Asp Gly Gly Ala Asp Thr Ala Thr Asp Asp Pro Cys Pro Cys  
130 135 140  
Ala

5

(2) INFORMATION FOR SEQ ID NO:207:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 525 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ix) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..525

(D) OTHER INFORMATION: / Ceres Seq. ID 1396782

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:207:

atgattttta cagacagag gagacggcga caatggcga ttgatgatg agataccac 60  
ttagcttgg agctttctt gttcagctt cttca-ccaa cggctgcgc cggatgatg 120  
gagctatag cggctgtga ggcctgga ttgtggaatt accgttgag aagatagaa 180  
gaucttgat cggacacaga tcaacatga agaacaagt gaagagctt atgatatga 240  
tcgcgaatt cggcttcaa gttccgattg acgtgattga agttagtga acttactatg 300  
ggcttggg atgtcacaga tccggggc accagaagt aggtttcca acaatagtt 360  
gcaaatcag taaggaaaca aagaaacat -aagaaatca tcttcgtga agagagttt 420  
actttatgt atattgttg tatgataga actcagatg gtatgtaa gttaagaata 480  
agaacttgt acaacttgt gataattta gtagctgat ttgc

(2) INFORMATION FOR SEQ ID NO:208:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 135 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..135

(D) OTHER INFORMATION: / Ceres Seq. ID 1396703

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:208:

Asp Phe Tyr Arg Arg Glu Glu Thr Ala Thr Met Ala Asn Leu Met Met 5  
1 Arg Leu Pro Ile Ser Leu Arg Ser Phe Ser Val Ser Ala Ser Ser Ser 10  
20 25 30  
Asn Gly Ser Pro Val Ile Gly Gly Ser Ser Gly Val Gly Pro 35  
35 40 45  
Met Ile Val Glu Leu Pro Leu Glu Lys Ile Arg Arg Pro Leu Met Arg 50  
50 55 60  
Thr Arg Ser Asn Asp Gln Asn Lys Val Lys Glu Leu Met Asp Ser Ile 65  
65 70 75 80  
Arg Gln Ile Gly Leu Gln Val Pro Ile Asp Val Ile Glu Val: Asp Gly 85  
90 95

50

Thr Tyr Tyr Gly Phe Ser Gly Cys His Arg Tyr Glu Ala His Gln Lys 100  
100 105 110 115  
Leu Gly Leu Pro Thr Ile Arg Cys Lys Ile Arg Lys Gly Thr Lys Glu 125  
125 130 135

55

(2) INFORMATION FOR SEQ ID NO:209:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 125 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

60

422

(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..125

(D) OTHER INFORMATION: / Ceres Seq. ID 1396784

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:209:

Met Ala Asn Leu Met Met Arg Leu Pro Ile Ser Leu Arg Ser Phe Ser 1  
1 5 10 15  
Val Ser Ala Ser Ser Asn Gly Ser Pro Pro Val Ile Gly Gly Ser 20  
20 25 30  
Ser Gly Gly Val Gly Pro Met Ile Val Glu Leu Pro Leu Glu Lys Ile 35  
35 40 45  
Arg Arg Pro Leu Met Arg Thr Arg Ser Asn Asp Gln Asn Lys Val Lys 50  
50 55 60  
Glu Leu Met Asp Ser Ile Arg Gln Ile Gly Leu Gln Val Pro Ile Asp 65  
65 70 75 80  
Val Ile Glu Val Asp Gly Thr Tyr Tyr Gly Phe Ser Gly Cys His Arg 85  
90 95  
Tyr Glu Ala His Gln Lys Leu Gly Leu Pro Thr Ile Arg Cys Lys Ile 100  
100 105 110 115  
Arg Lys Gly Thr Lys Glu Thr Leu Arg His His Leu Arg 120  
120 125

5

10

15

20

25

30

35

40

45

50

55

60

(1)

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 939 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..939

(D) OTHER INFORMATION: / Ceres Seq. ID 1396802

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:211:

423  
atacaaac aggtcaggg tttatcttc catattctt tccactgat tttacaatc  
60 gattcaagg aattacgaa ctgtttgta attcttggt gaaagagct tgaattgct  
120 gtttgtttg gattcaagt gactgtgtg ttgtttgtg gtaacagct gatgactct  
180 tcaagacat catgttttg tctgtatct cacagagcg gtaccatag gaggaagaag  
240 attcttgga aagaagaaa tctgaatta ggaagctct tcaattgct acyattgtt  
300 ttatgcttc ctcgaagag tctatcgaga attcgcaat gaaagagcaa aagctcttg  
360 attgtaatt agaatgtgc tgaattatc gactatact tctgtatgt acaaaagcaa  
420 attgtaatt atttatac caaagtctg agaacgttc ttaactagt gaagaaatg  
480 aactcttc attacactg tttatataa catatcataa caaacgacc gactgagc  
540 aactcttc tctgtgctt cgcgaagag agtcaggag ttccagaag agtatgac  
600 acgcttct actgtatct caatgagtc agaaatgag accatggaa gatacaca  
660 gaaatagg argaaactt ggcgttgat cgtgaagcc tgaaggagc cgtgatag  
720 ccttcagtg acacaactt aagtaacag cgtgaagag acagcgaca caacgaca  
780 acattgct tgaatttg tttcaagt tttagactt cgaattttt agttgatta  
840 catattacta tsgaagaa gactgtgta tatgtctt attcttgga tgaactat  
900 gatataagg gcaaaaaca acactctct tctgtctct  
(2) INFORMATION FOR SEQ ID NO:212:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 248 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..248  
(D) OTHER INFORMATION: / Ceres Seq. ID 1396803  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:212:  
Ile Thr Asn Gln Gly Gln Gly Phe Ile Phe His Ile Leu Phe Pro Leu  
1 10  
Ile Leu Thr Ile Asp Leu Arg Glu Ser Thr Asn Cys Leu Val Ile Ser  
20 30  
Gly Gly Lys Ser Leu Asp Leu Ala Val Trp Phe Gly Val Met Val Ser  
35 45  
Ala Trp Phe Val Val Gly Ser Pro Val Met Asp Ser Ser Ser Pro  
50 60  
Cys Leu Cys Leu Asp Ala His Thr Thr Gly Thr Ile Arg Arg Lys Lys  
65 75  
Ile Leu Gly Lys Ala Arg Asn Leu Glu Leu Gly Ser Ser Phe Thr Gly  
85 95  
Ser Arg Ile Val Phe Arg Leu Ser Pro Lys Arg Val Ser Arg Ile Ala  
100 110  
Asn Arg Lys Ser Lys Lys Leu Leu Ile Val Asn Glu Asp Val Ala Gly  
115 125  
Asn Tyr Asp Asp Thr Phe Gly Asp Val Gln Lys Gln Ile Val Asn Tyr  
130 140  
Phe Thr Tyr Lys Ala Val Arg Thr Val Leu Leu Phe Gln Leu Tyr Glu Met  
145 155  
Asn Pro Pro Gln Tyr Thr Trp Phe Tyr Asn His Ile Ile Thr Asn Arg  
160 175  
Pro Thr Asp Gly Lys Arg Phe Leu Arg Ala Leu Gly Lys Glu Ser Gln  
180 195  
Glu Leu Ala Glu Arg Val Met Ile Thr Arg Leu His Leu Tyr Gly Lys  
200 210  
Trp Ile Lys Lys Cys Asp His Gly Lys Ile Tyr Gln Glu Ile Ser Asp  
215 225  
Glu Asn Leu Ala Leu Met Arg Glu Arg Leu Met Glu Thr Val Ile Trp  
230 240  
Pro Ser Asp Thr Asn Ser Arg  
245  
(2) INFORMATION FOR SEQ ID NO:213:  
(i) SEQUENCE CHARACTERISTICS:

424  
(A) LENGTH: 203 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..203  
(D) OTHER INFORMATION: / Ceres Seq. ID 1396804  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:213:  
Met Val Ser Ala Trp Phe Val Val Gly Ser Pro Val Met Asp Ser Ser  
1 10  
Ser Ser Pro Cys Leu Cys Leu Asp Ala His Thr Thr Gly Thr Ile Arg  
20 30  
Arg Lys Lys Ile Leu Gly Lys Ala Arg Asn Leu Glu Leu Gly Ser Ser  
35 45  
Phe Thr Gly Ser Arg Ile Val Phe Arg Leu Ser Pro Lys Arg Val Ser  
50 60  
Arg Ile Ala Asn Arg Lys Ser Lys Lys Leu Leu Ile Val Asn Glu Asp  
65 75  
Val Ala Gly Asn Tyr Asp Asp Thr Phe Gly Asp Val Gln Lys Gln Ile  
85 95  
Val Asn Tyr Phe Thr Tyr Lys Ala Val Arg Thr Val Leu His Gln Leu  
100 110  
Tyr Glu Met Asn Pro Pro Gln Tyr Thr Trp Phe Tyr Asn His Ile Ile  
115 125  
Thr Asn Arg Pro Thr Asp Gly Lys Arg Phe Leu Arg Ala Leu Gly Lys  
130 140  
Glu Ser Gln Glu Leu Ala Glu Arg Val Met Ile Thr Arg Leu His Leu  
145 155  
Tyr Gly Lys Trp Ile Lys Lys Cys Asp His Gly Lys Ile Tyr Gln Glu  
160 175  
Ile Ser Asp Glu Asn Ala Leu Met Arg Glu Arg Leu Met Glu Thr  
180 195  
Val Ile Trp Pro Ser Asp Asp Thr Asn Ser Arg  
200 210  
(2) INFORMATION FOR SEQ ID NO:214:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 191 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..191  
(D) OTHER INFORMATION: / Ceres Seq. ID 1396805  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:214:  
Met Asp Ser Ser Ser Pro Cys Leu Cys Leu Asp Ala His Thr Thr  
1 10  
Gly Thr Ile Arg Arg Lys Lys Ile Leu Gly Lys Ala Arg Asn Leu Glu  
15 20  
Leu Gly Ser Ser Phe Thr Gly Ser Arg Ile Val Phe Arg Leu Ser Pro  
25 30  
Lys Arg Val Ser Arg Ile Ala Asn Arg Lys Ser Lys Leu Leu Ile  
35 45  
Val Asn Glu Asp Val Ala Gly Asn Tyr Asp Asp Thr Phe Gly Asp Val  
50 60  
Gln Lys Gln Ile Val Asn Tyr Phe Thr Tyr Lys Ala Val Arg Thr Val  
65 75  
Leu His Gln Leu Tyr Glu Met Asn Pro Pro Gln Tyr Thr Trp Phe Tyr  
80 95  
100 110

425  
 115 Thr Tyr Tyr Phe Thr Pro Thr Asp Gly Lys Arg Phe Leu Arg  
 120 Ala Leu Gly Lys Glu Ser Gln Glu Leu Ala Glu Arg Val Met Ile Thr  
 130 Arg Leu His Leu Tyr Gly Lys Trp Ile Lys Lys Cys Asp His Gly Lys  
 145 Ile Tyr Gln Glu Ile Ser Asp Glu Asn Leu Ala Leu Met Arg Glu Arg  
 165 Leu Met Glu Thr Val Ile Trp Pro Ser Asp Thr Asn Ser Arg  
 180 Ser Met Glu Thr Val Ile Trp Pro Ser Asp Thr Asn Ser Arg  
 185  
 190  
 (2) INFORMATION FOR SEQ ID NO:215:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 753 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..753  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1397130  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:215:  
 aaagattgga agaaaagt tactaaaaa gaacccgggg aagtaagta gataagagaa  
 gagagagaga tcggaagacc agcgaacaa aagaccagga aagatgtgc tctggagag  
 ttgattgata cctatgtcgc acagtcgcat aactgcaca agtgagaggt gattgatac  
 caggaggaat atgaagatac cagaagtaaa atgtctgagg atccttttaa ctgcagaa  
 aaacaggaga tctcttgaga agacccgct gatactgact agcatcttc tggacttgg  
 gcatctgaca agctctgctt ccccaaacg cctaaaggtt tcaagaaag attagttctc  
 agaaagatt acttaagat ggatactac tacttaact ctaacgggaa agacttcag  
 actccaatt aaacccctgc ctctgtgaa ccccaaccgg aattcagaa cgaaccactt  
 gagactta atttactgt cccaagctc atggaagata ctgtccccc tgaaccgaag  
 ctgtctctc atttctcaag cactcaccac actactccag agaaagcag tgcagagag  
 agcctaact aatatgtct ctctccggg agttctteta ctctcttct gctccctaaa  
 gaattgggt ttgagcttt atagatttga tttcttctaa aactgttatg atataact  
 cctaaagca aaaaataga tctttgtct etc  
 (2) INFORMATION FOR SEQ ID NO:216:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 203 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..203  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1397131  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:216:  
 Lys Asp Leu Glu Lys Lys Phe Thr Lys Lys Lys Glu Ser Gly Glu Val Ser  
 1 5 10 15  
 Glu Met Lys Glu Glu Glu Ile Gly Lys Pro Ala Lys Pro Lys Ala  
 20 25 30  
 Lys Lys Asp Val Ala Pro Gly Arg Leu Ile Asp Thr Tyr Ala Ala Gln  
 35 40 45  
 Cys Asp Asn Cys His Lys Trp Arg Val Ile Asp Ser Gln Glu Glu Tyr  
 50 55 60  
 Glu Asp Ile Arg Ser Lys Met Leu Glu Asp Pro Phe Asn Cys Gln Lys  
 65 70 75 80  
 Lys Gln Gly Met Ser Cys Glu Glu Pro Ala Asp Ile Asp Tyr Asp Ser  
 85 90 95  
 Ser Arg Thr Trp Val Ile Asp Lys Pro Gly Leu Pro Lys Thr Pro Lys  
 100 105 110  
 Gly Phe Lys Arg Ser Leu Val Leu Arg Lys Asp Tyr Ser Lys Met Asp

426  
 115 Thr Tyr Tyr Phe Thr Pro Thr Thr Gly Lys Leu Arg Ser Arg Asn Glu  
 130 Ile Ala Ala Phe Val Glu Ala Asn Pro Glu Phe Arg Asn Ala Pro Leu  
 145 Gly Asp Phe Asn Phe Thr Val Pro Lys Val Met Glu Asp Thr Val Pro  
 165 Pro Asp Pro Lys Leu Gly Ser Pro Phe Pro Ser Thr Thr Thr Thr  
 180 Ser Glu Lys Ser Ser Val Lys Gln Ser His Asn  
 195  
 200  
 (2) INFORMATION FOR SEQ ID NO:217:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 186 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..186  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1397132  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:217:  
 Met Lys Glu Glu Glu Glu Ile Gly Lys Pro Ala Lys Pro Lys Ala Lys  
 1 5 10 15  
 Lys Asp Val Ala Pro Gly Arg Leu Ile Asp Thr Tyr Ala Ala Gln Cys  
 20 25 30  
 Asp Asn Cys His Lys Trp Arg Val Ile Asp Ser Gln Glu Glu Tyr Glu  
 35 40 45  
 Asp Ile Arg Ser Lys Met Leu Glu Asp Pro Phe Asn Cys Gln Lys Lys  
 50 55 60  
 Gln Gly Met Ser Cys Glu Glu Pro Ala Asp Ile Asp Tyr Asp Ser Ser  
 65 70 75 80  
 Arg Thr Trp Val Ile Asp Lys Pro Gly Leu Pro Lys Thr Pro Lys Gly  
 85 90 95  
 Phe Lys Arg Ser Leu Val Leu Arg Lys Asp Tyr Ser Lys Met Asp Thr  
 100 105 110  
 Tyr Tyr Phe Thr Thr Thr Thr Gly Lys Lys Arg Ser Arg Asn Glu Ile  
 115 120 125  
 Ala Ala Phe Val Glu Ala Asn Pro Glu Phe Arg Asn Ala Pro Leu Gly  
 130 135 140  
 Asp Phe Asn Phe Thr Val Pro Lys Val Met Glu Asp Thr Val Pro Pro  
 145 150 155  
 Asp Pro Lys Leu Gly Ser Pro Phe Pro Ser Thr Thr Thr Thr Ser  
 165 170 175  
 Glu Lys Ser Ser Val Lys Gln Ser His Asn  
 180 185 190  
 (2) INFORMATION FOR SEQ ID NO:218:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 133 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..133  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1397133  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:218:  
 Met Leu Glu Asp Pro Phe Asn Cys Gln Lys Lys Gln Gly Met Ser Cys  
 1 5 10 15  
 Gln Glu Pro Ala Asp Ile Asp Tyr Asp Ser Ser Arg Thr Trp Val Ile



427

20 25 30  
 Asp Lys Pro Gly Leu Pro Lys Thr Pro Lys Gly Phe Lys Arg Ser Leu  
 35 40 45  
 Val Leu Arg Lys Asp Tyr Ser Lys Met Asp Thr Tyr Thr Phe Thr Pro  
 50 55 60  
 Thr Gly Lys Lys Leu Arg Ser Arg Asn Glu Ile Ala Phe Val Glu  
 65 70 75 80  
 Ala Asn Pro Glu Phe Arg Asn Ala Pro Leu Gly Asp Phe Asn Phe Thr  
 85 90 95  
 Val Pro Lys Val Met Glu Asp Thr Val Pro Pro Asp Pro Lys Leu Gly  
 100 105 110  
 Ser Pro Phe Pro Ser Thr Thr Thr Thr Ser Glu Lys Ser Ser Val  
 115 120 125  
 Lys Gln Ser His Asn

(2) INFORMATION FOR SEQ ID NO:219:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 784 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..784

(D) OTHER INFORMATION: / Ceres Seq. ID :398004

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:219:

actcaccac caaacacaa cataaaacaa caagtggaag ctttaaaac agagggagag  
 60  
 agcaaaaag ggcagctcg gacgctacg gacggaagt ccgctaaag gacggcgga  
 120  
 gaatactac aagagtgga agaacagaa ccatgttc cc-gatgcta tcggcacca  
 180  
 catccaagt gttacattc acgaagga acataact caggtgtca tcagagttg  
 240  
 gaactaca tggatgga agggaggt g-tcaaggag agaaagaga tagagatga  
 300  
 gacaaagc tggcgtta gggatgtga gggtcnctg atggagcgc tcaaggtga  
 360  
 cgactgtc taccattca tcccaattc tggatgacc tgcctcgga aatcacttt  
 420  
 aataggca agatgcacg agtatccc agaaccaagc ggtcacataa aatgcgtca  
 480  
 gactgtgt gtcacatg gaacacgt tagcaaat taactcat tcccacgtc  
 540  
 gtccgtcg tcatcatc caccatcc atcatatca tcatcatc caccatcc  
 600  
 atcatatca ctatctgat ttaagtta agatgttc agtataaa atgggtctt  
 660  
 gtggatcgt cattctatg tgaacact tggttctg atgatgtc gatattgt  
 720  
 tatgttcag atcatatgc ggggtcgata taagtatct taagttaatt ttactacaa  
 780  
 ttcc

(2) INFORMATION FOR SEQ ID NO:220:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 151 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..151

(D) OTHER INFORMATION: / Ceres Seq. ID :1398005

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:220:

Met Ala Thr Ser Gly Thr Tyr Val Thr Glu Val Pro Leu Lys Gly Ser  
 1 5 10 15  
 Ala Glu Lys Tyr Tyr Lys Arg Trp Lys Asn Glu Asn His Val Phe Pro  
 20 25 30  
 Asp Ala Ile Gly His His Ile Gln Asn Val Thr Val His Glu Gly Glu  
 35 40 45  
 His Asp Ser His Gly Ser Ile Arg Ser Trp Asn Tyr Thr Trp Asp Gly  
 50 55 60  
 Lys Glu Glu Val Phe Lys Glu Arg Arg Glu Ile Asp Asp Glu Thr Lys

428

65 70 75 80  
 Thr Leu Thr Leu Arg Gly Leu Glu Gly His Val Met Glu Gln Leu Lys  
 85 90 95  
 Val Tyr Asp Val Val Tyr Gln Phe Ile Pro Lys Ser Glu Asp Thr Cys  
 100 105 110  
 Ile Gly Lys Ile Thr Leu Ile Trp Glu Lys Arg Asn Asp Asp Ser Pro  
 115 120 125  
 Glu Pro Ser Gly Tyr Ile Lys Phe Val Lys Ser Leu Val Ala Asp Met  
 130 135 140  
 Gly Asn His Val Ser Lys Thr  
 145 150  
 Ser Pro Phe Pro Ser Thr Thr Thr Thr Ser Glu Lys Ser Ser Val

(2) INFORMATION FOR SEQ ID NO:221:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 389 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..389

(D) OTHER INFORMATION: / Ceres Seq. ID :1398370

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:221:

aaaaacttt tcagagtcgg aggtttcaa caaqaqaa gaacatgaag cgaacgtga  
 60  
 tggacaagc gattcatca tgggtgtg tgggtgtg attggttt ggtacttgt  
 120  
 cactggaact tgggtacaag cttctcttg aaaggtga acaatcgaa agatctctc  
 180  
 agttcttca acaatcaaa caacaagatg acaagaaga agcaggtgg gacaataga  
 240  
 atgcagggg gtggagagag aagagtag cttaaatcg gttcaatc agacatcca  
 300  
 aagtgtaa ggaagaga cttatcaaa attttctt taatcaacg taagtcaat  
 360  
 ctgttaacta tcaacatgg cttctttt

(2) INFORMATION FOR SEQ ID NO:222:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 88 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..88

(D) OTHER INFORMATION: / Ceres Seq. ID :1398371

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:222:

Lys Leu Phe Gln Ser Arg Arg Phe Gln Thr Glu Arg Lys Lys Met Lys  
 1 5 10 15  
 Arg Thr Val Met Asp Asn Ala Ile Arg Ser Ser Val Val Val Leu Gly  
 20 25 30  
 Ser Leu Ala Phe Tyr Tyr Leu Ser Leu Glu Leu Gly Tyr Lys Pro Phe  
 35 40 45  
 Leu Glu Lys Ala Glu Gln Tyr Glu Arg Ser Leu Gln Ser Ser Gln Gln  
 50 55 60  
 His Gln Gln Gln Asp Glu Gln Glu Ala Arg Trp Asp Asn Ser Asn  
 65 70 75 80  
 Val Glu Gly Trp Glu Glu Lys Arg

(2) INFORMATION FOR SEQ ID NO:223:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 74 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

429

(A) NAME/KEY: peptide  
(B) LOCATION: 1..74

(D) OTHER INFORMATION: / Ceres Seq. ID 1399372

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:223:

5 Met Lys Arg Thr Val Met Asp Asn Ala Ile Arg Ser Ser Val Val Val  
1 10  
Leu Gly Ser Leu Ala Phe Gly Tyr Leu Ser Leu Glu Leu Gly Tyr Lys  
20 30  
Pro Phe Leu Glu Lys Ala Glu Gln Tyr Glu Arg Ser Leu Gln Ser Ser  
35 45  
Gln Gln His Gln Gln Asp Glu Gln Gln Glu Ala Arg Trp Asp Asn  
50 60  
Ser Asn Val Glu Gly Trp Glu Glu Lys Arg  
65 70

(2) INFORMATION FOR SEQ ID NO:224:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 69 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..69

(D) OTHER INFORMATION: / Ceres Seq. ID 1399373

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:224:

25 Met Asp Asn Ala Ile Arg Ser Ser Val Val Val Leu Gly Ser Leu Ala  
1 10  
Phe Gly Tyr Leu Ser Leu Glu Leu Gly Tyr Lys Pro Phe Leu Glu Lys  
20 30  
Ala Glu Gln Tyr Glu Arg Ser Leu Gln Ser Ser Gln Gln His Gln Gln  
35 45  
Gln Asp Glu Gln Glu Ala Ala Arg Trp Asp Asn Ser Asn Val Glu Gly  
50 60  
Trp Glu Glu Lys Arg  
65 70

(2) INFORMATION FOR SEQ ID NO:225:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 745 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..745

(D) OTHER INFORMATION: / Ceres Seq. ID 1425147

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:225:

50 atttctctgt cctctgtctc aaactcgaat caaaatctct aaagtctcaa ttttttgtt  
120 cgttctttt tttttttta aagaatggct tcaattctgt caacttgcg ttgcctatg  
180 ttaactcac agagaatac caatgcaca tgaatcaca aactctcat tttcttaact  
240 caggttaca agtatcttc tcaattctt cttactcga accatcgag catcagctg  
300 atggtgaat ttgttgataa gttctgagat tcatcgactg atctcgactg ttttagtgc  
360 attcagaat ttgttgataa gttctgagat aggttagctc ttatcggtt gggtttgct  
420 ggtatcgag cctctggcg atcattgat cctcatcagg caatgcaca attgcctgt  
480 attctgagc gattcgact agttgtatc ttgttctca cgtgttcat atctcgtat  
540 cctctgtca aacgcagag acgagagct ttgaaattg tcaagaatc agtagcgat  
600 atactggcc gttgagact gttgttgcg taatactca ttcttggaag atgattgtt  
660 tgcagtttg taatactca tgcagagtg gttattgtt ttatgcacaa aatgcctgc  
720 attgaacac tgaatacat ttatcttgg tttttgttg tgaacaaa caatcttcc  
taattcaaa gattctctt taagt

(2) INFORMATION FOR SEQ ID NO:226:

430

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 156 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..156

(D) OTHER INFORMATION: / Ceres Seq. ID 1425148

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:226:

10 Met Ala Ser Ile Ser Ala Thr Leu Pro Ser Pro Leu Leu Leu Thr Gln  
1 10  
Arg Lys Ser Asn Leu Thr Ser Ile Gln Lys Leu Pro Phe Ser Leu Thr  
15 20  
Arg Gly Thr Asn Asp Leu Ser Pro Leu Ser Leu Thr Arg Asn Pro Ser  
35 40  
Ser Ile Ser Leu Met Val Lys Ser Ser Gly Glu Ser Ser Asp Ser Ser  
50 55  
Thr Asp Leu Asp Val Val Ser Thr Ile Gln Asn Val Trp Asp Lys Ser  
65 70  
Glu Asp Arg Leu Gly Leu Ile Gly Leu Gly Phe Ala Gly Ile Val Ala  
85 90  
Leu Trp Ala Ser Leu Asn Leu Ile Thr Ala Ile Asp Lys Leu Pro Val  
100 105  
Ile Ser Ser G-y Phe Glu Leu Val Gly Ile Leu Phe Ser Thr Trp Phe  
115 120  
Thr Tyr Arg Tyr Leu Leu Phe Lys Pro Asp Arg Gln Glu Leu Ser Lys  
130 135  
Ile Val Lys Lys Ser Val Ala Asp Ile Leu Gly Gln  
145 150  
(2) INFORMATION FOR SEQ ID NO:227:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 104 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..104

(D) OTHER INFORMATION: / Ceres Seq. ID 1425149

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:227:

45 Met Val Lys Ser Ser Gly Glu Ser Ser Asp Ser Ser Thr Asp Leu Asp  
1 5  
Val Val Ser Thr Ile Gln Asn Val Trp Asp Lys Ser Glu Asp Arg Leu  
20 25  
Gly Leu Ile Gly Leu Gly Phe Ala Gly Ile Val Ala Leu Trp Ala Ser  
35 40  
Leu Asn Leu Ile Thr Ala Ile Asp Lys Leu Pro Val Ile Ser Ser Gly  
50 55  
Phe Glu Leu Val Gly Ile Leu Phe Ser Thr Thr Trp Phe Thr Arg Tyr  
65 70  
Leu Leu Phe Lys Pro Asp Arg Gln Glu Leu Ser Lys Ile Val Lys Lys  
85 90  
Ser Val Ala Asp Ile Leu Gly Gln  
100 105

(2) INFORMATION FOR SEQ ID NO:228:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 763 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single



433  
Ser Glu Ala Thr Glu Ala Lys Glu Ala Val Ala Ile Lys Glu Ala Val  
145 150 155 160  
Ala Val Lys Glu Ala Ala

5 (2) INFORMATION FOR SEQ ID NO:232:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 607 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(11) MOLECULE TYPE: DNA (genomic)  
(1x) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..607

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:232:

60  
aaactccta aagttcaat ttttktgtc tgtctctt tttttaag aatgcttca  
120  
attcgcga cttgcctt gccattgta ctacacaga gaatacaca tctcacatg  
180  
attcaaac tccatttt tctactcga gttacgaag acctctcc attactctt  
240  
actgaaac ctacagcat cactcgtat gttaaacta gtgagaaag ctacatcca  
300  
tcgaccgatc tgcagctgt cagtaagat cagaatgta caatgaca atgcgcgt  
360  
atctcgatg gattgcagt agttgatac ttgtctcca gttgtctac atctgcat  
420  
ctctcttcca aaccgaaag acaagagct tcaaaatgt tcaagaatc agttgagat  
480  
atactggcc agtgaacct gttgtgtga taactacca tcttgaaag atgatttgt  
540  
tgcagctgt taaattcca tgaacaggt gttgttgtt ctactcaat atgtcagc  
600  
attgaaac tgaataact ttaactgtg ttttgttg tgaacaaat caactcttc  
taatttc

30 (2) INFORMATION FOR SEQ ID NO:233:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 127 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(11) MOLECULE TYPE: peptide  
(1x) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..127

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:233:

15  
Met Ala Ser Ile Ser Ala Thr Leu Pro Ser Pro Leu Leu Thr Thr  
19  
Arg Lys Ser Asn Leu Thr Ser Ile Gln Lys Leu Pro Phe Ser Leu Thr  
25  
Arg Gly Thr Asn Asp Leu Ser Pro Leu Ser Leu Thr Arg Asn Pro Ser  
35  
Ser Ile Ser Leu Met Val Lys Ala Ser Gly Glu Ser Ser Asp Ser Ser  
45  
Thr Asp Leu Asp Val Val Ser Thr Ile Gln Asn Val Ala Ile Asp Lys  
55  
Leu Pro Val Ile Ser Ser Gly Phe Glu Leu Val Gly Ile Leu Phe Ser  
65  
Thr Thr Phe Thr Tyr Arg Tyr Leu Leu Phe Lys Pro Asp Arg Gln Glu  
75  
Leu Ser Lys Ile Val Lys Lys Ser Val Ala Asp Ile Leu Gly Gln  
110  
115 120 125

- (2) INFORMATION FOR SEQ ID NO:234:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 75 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(11) MOLECULE TYPE: peptide

434  
(1x) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..75  
(D) OTHER INFORMATION: / Ceres Seq. ID 1447482

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:234:

1 Met Val Lys Ala Ser Gly Glu Ser Ser Asp Ser Thr Asp Leu Asp  
5  
Val Val Ser Thr Ile Gln Asn Val Ala Ile Asp Lys Leu Pro Val Ile  
10  
Ser Ser Gly Phe Glu Leu Val Gly Ile Leu Phe Ser Thr Thr Phe Thr  
15  
Tyr Arg Tyr Leu Leu Phe Lys Pro Asp Arg Gln Glu Leu Ser Lys Ile  
20  
Val Lys Lys Ser Val Ala Asp Ile Leu Gly Gln  
25  
65 70 75

15 (2) INFORMATION FOR SEQ ID NO:235:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 668 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(11) MOLECULE TYPE: DNA (genomic)  
(1x) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..668

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:235:

60  
aanaaaac aaacaaaa attattcca agagaagag gaanaatga attctctc  
120  
cgaatgga aagaacac caagctcac accagagcc agacacac aagctcag  
180  
aagaaacga aagaatcca agaccgcta ccaacgcga gtccatgga atgtcagag  
240  
ttgagctg agctgtcca gccagcagc gtaacgaac agacaatc gcaaaagta  
300  
aagctgcg agctctct gatactag agctgcga gaataacgt agttcgatg  
360  
aagaatgt cactgtcag taactcga agttgata gttctaac gactcaagt  
420  
gcacacac cactgtct gttgtctc ctcccgac gactcagc gactcagc  
480  
gtcagctc gccagagct aagaagag atgaagtc tggctgag ctgtgaggt  
540  
atgcagat gttcagat tcttgagat gattgact ttaattgt ttcaatct  
600  
tgaataat aaatcaata actagatc ttgtgact gttatgtg ctgccttat  
660  
gttaaggg agtgaagt gattgata attcctgt atcatgac taactact  
tcgtgtc

40 (2) INFORMATION FOR SEQ ID NO:236:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 62 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(11) MOLECULE TYPE: peptide  
(1x) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..62

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:236:

15  
Lys Lys Asn Lys Thr Lys Asn Tyr Ile Gln Glu Lys Lys Glu Lys Met  
10  
Asn Phe Ile Ser Asp Gln Val Lys Lys Ser Ser Ser Thr Pro Glu  
20  
Ala Arg Pro Gln Gln Ala Ser Arg Arg Asn Arg Asn Ser Tyr Lys Thr  
25  
Ser Tyr Gln Arg Arg Ala His Gly Lys Cys Gln Gly Cys Ser  
30  
60 65 70

- (2) INFORMATION FOR SEQ ID NO:237:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 115 amino acids

435

(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..115  
(D) OTHER INFORMATION: / Ceres Seq. ID 1447579  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:237:  
1 Met Ala Ser Ala Lys Val Val Ala Glu Ala Ala Glu Ala Ala Ala 15  
Asn Glu Ser Asp Lys Leu Asp Lys Gly Lys Val Ala Gly Ala Ser Ala 30  
Asp Ile Leu Asp Ala Ala Glu Lys Tyr Gly Lys Phe Asp Glu Lys Ser 45  
Ser Thr Gly Gln Tyr Leu Asp Lys Ala Glu Lys Tyr Leu Asn Asp Tyr 60  
Glu Ser Ser His Ser Thr Gly Ala Gly Gly Pro Pro Pro Thr Ser 75  
65 Glu Ala Glu Pro Ala Ser Gln Pro Glu Pro Ala Ala Lys Lys Asp Asp 80  
Glu Glu Ser Gly Gly Gly Leu Gly Gly Tyr Ala Lys Met Ala Gln Gly 95  
Phe Leu Lys 100  
115  
(2) INFORMATION FOR SEQ ID NO:238:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 69 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..69  
(D) OTHER INFORMATION: / Ceres Seq. ID 1447580  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:238:  
1 Met Lys Arg Val Ala Leu Val Ser Thr Ser Thr Arg Leu Arg Ser Ile 15  
Ser Thr Thr Thr Ser Arg His Thr Pro Pro Val Leu Val Val Leu Leu 30  
Leu Arg Arg Val Arg Leu Ser Gln Gln Val Ser Leu Ser Arg Arg Leu 45  
Arg Lys Thr Met Lys Ser Leu Val Val Gly Leu Glu Val Met Pro Arg 60  
Trp Leu Lys Val Ser 65  
(2) INFORMATION FOR SEQ ID NO:239:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 717 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ix) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..717  
(D) OTHER INFORMATION: / Ceres Seq. ID 1447922  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:239:  
60 ctctgtctcgttgcctcaaaaacgaatcaaaatctctcaaaagtttcaattttttttgttctgt  
tattttttttttttaaagaatgggttcaaaatltctgcaac ttgtctctg ccattgtac  
tcaacagag aaaaatccaat ctacataga ttcaaaaact ccaattttttt ctaactcagag 120  
130

436

gracgaatca tttttctcca ttactcttca ctggaaccc tagcagcatc agtctgatgg 240  
tgaagctag tggagaagac tcagattcat cgaactgact cgaactgttt agtcagattc 300  
agaatggtt ggaatagctt gaagataggt taggtctat ttgtttggtt ttgttggtta 360  
tttagcttt ttgggcatca ttgaatcca tcacggcaat tgacaattg cccgttatct 420  
cgagcgatt cgaatagttt ggtattttgt tctccactg gttccatat cgaatatct 480  
tttcaaac ggaacgagag gaggcttcca aaatttcca aaatcagta gggatatac 540  
ttggcagtg acccttgtt ggtgataat acttcatct tgaagaatg ttgtttgca 600  
agttgaaa attacattac aggtggttt ttgttccag tcaataatg tcatgattt 660  
(2) INFORMATION FOR SEQ ID NO:240:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 156 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..156  
(D) OTHER INFORMATION: / Ceres Seq. ID 1447923  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:240:  
1 Met Ala Ser Ile Ser Ala Thr Leu Pro Ser Pro Leu Leu Leu Thr Gln 15  
Arg Lys Ser Asn Leu Thr Ser Ile Gln Lys Leu Pro Phe Ser Leu Thr 30  
Arg Gly Thr Asn Asp Leu Ser Pro Leu Thr Arg Asn Pro Ser 45  
Ser Ile Ser Leu Met Val Lys Ala Ser Gly Glu Ser Ser Asp Ser Ser 60  
Thr Asp Leu Asp Val Val Ser Thr Ile Gln Asn Val Trp Asp Lys Ser 75  
Glu Asp Arg Leu Gly Leu Ile Gly Leu Gly Phe Ala Gly Ile Val Ala 90  
Leu Trp Ala Ser Leu Asn Leu Ile Thr Ala Ile Asp Lys Leu Pro Val 110  
Ile Ser Ser Gly Phe Glu Leu Val Gly Ile Leu Phe Ser Thr Trp Phe 125  
Thr Tyr Arg Tyr Leu Leu Phe Lys Pro Asp Arg Gln Glu Leu Ser Lys 140  
Ile Val Lys Lys Ser Val Ala Asp Ile Leu Gly Gln 155  
(2) INFORMATION FOR SEQ ID NO:241:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 104 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..104  
(D) OTHER INFORMATION: / Ceres Seq. ID 1447924  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:241:  
1 Met Val Lys Ala Ser Gly Glu Ser Ser Ser Asp Ser Ser Thr Asp Leu Asp 15  
Val Val Ser Thr Ile Gln Asn Val Trp Asp Lys Ser Glu Asp Arg Leu 30  
Gly Leu Ile Gly Leu Gly Phe Ala Gly Ile Val Ala Leu Trp Ala Ser 45  
Leu Asn Leu Ile Thr Ala Ile Asp Lys Leu Pro Val Ile Ser Ser Gly 60  
Phe Glu Leu Val Gly Ile Leu Phe Ser Thr Trp Phe Thr Tyr Arg Tyr 80

437

438

65 Leu Leu Phe Lys Pro Asp Arg Gln Gln Leu Ser Lys Ile Val Lys Lys 80  
85 Ser Val Ala Asp Ile Leu Gly Gln 90  
100

5 (2) INFORMATION FOR SEQ ID NO:242:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 656 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..656  
(D) OTHER INFORMATION: / Ceres Seq. ID 1448012  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:242:  
ggattctc accctccta cctctcttag ccagctctga cctctcgat tctttccgg  
tgaagaggga gcagaaacat ggtgtcaggt accagattc cgagattggg atgcacaaat  
cgccattt tttaggttat ggtgtgat agcagattc caagaga-gg gaagcattt  
gaagcttag gttacttcaa ccttttcca guccagagcg gtgtaagag gagggtctc  
aaqctcac gaattaaaga ttgttact gttggctc agccatcaga cccgttcaa  
cgtctctt tcagatcgg ttacttct cctctcaa tgggtgctat ggaqctaaa  
ggtggagac ggcacacag cccagttgat ccaatgtagt gtcgtatgt gaatcgag  
aataaaacg ttaatgcga tgaatacag ccaagagag aggtctcaga agcaagatt  
gcagattca taagcttct gtgtgtag ctcttcagt tcaatttct gtcgattata  
ttgtcaag cagctcaga caactact gttctctt ttggcgata aacggcaag  
ttgttgac ttgttgaga caagcact atttgcatt aggatattt aatttt

10 Met Ala Ala Asp Ser Arg Ser Pro Arg Asp Gly Lys Lys Leu Glu Val  
Leu Gly Tyr Phe Asn Pro Leu Pro Gly Gln Asp Gly Gly Lys Arg Met  
Gly Leu Lys Phe Asp Arg Ile Lys Tyr Trp Leu Ser Val Gly Ala Gln  
15 Pro Ser Asp Pro Val Gln Arg Leu Leu Phe Arg Ser Gly Leu Pro  
Pro Pro Met Val Ala Met Gly Arg Lys Gly Gly Ala Arg Asp Thr  
Arg Pro Val Asp Pro Met Thr Gly Arg Tyr Val Asp Ala Glu Asn Lys  
Thr Val Asn Ala Asn Asp Asn Gln Pro Lys Glu Glu Asp Ser Glu Asp  
Lys Ile Ala  
115

20 (2) INFORMATION FOR SEQ ID NO:245:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 84 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..84  
(D) OTHER INFORMATION: / Ceres Seq. ID 1448015  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:245:  
Met Gly Leu Lys Phe Asp Arg Ile Lys Tyr Trp Leu Ser Val Gly Ala  
Gln Pro Ser Asp Pro Val Gln Arg Leu Leu Phe Arg Ser Gly Leu  
Pro Pro Pro Met Val Ala Met Gly Arg Lys Gly Gly Ala Arg Asp  
Thr Arg Pro Val Asp Pro Met Thr Gly Arg Tyr Val Asp Ala Glu Asn  
Lys Thr Val Asn Ala Asn Asp Asn Gln Pro Lys Glu Glu Asp Ser Glu  
Asp Lys Ile Ala  
70

25 (2) INFORMATION FOR SEQ ID NO:246:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 459 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..459  
(D) OTHER INFORMATION: / Ceres Seq. ID 1448135  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:246:  
aattctcag ggtttgaag agtgcctc gcgcggtg taattctct gtagcaaat  
cgacaaatg ggtcactcta atgtatggaa cttctatcgc aagaatcag gtcctggatc

30 (2) INFORMATION FOR SEQ ID NO:243:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 135 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..135  
(D) OTHER INFORMATION: / Ceres Seq. ID 1448013  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:243:  
Met Val Val Arg Ile Arg Leu Ser Arg Phe Gly Cys Lys Asn Arg Pro  
Phe Arg Val Met Ala Ala Asp Ser Arg Ser Pro Arg Asp Gly Lys  
His Leu Glu Val Leu Gly Tyr Phe Asn Pro Leu Pro Gly Gln Asp Gly  
Gly Lys Arg Met Gly Leu Lys Phe Asp Arg Ile Lys Tyr Trp Leu Ser  
Val Gly Ala Gln Pro Ser Asp Pro Val Gln Arg Leu Leu Phe Arg Ser  
Gly Leu Leu Pro Pro Pro Met Val Ala Met Gly Arg Lys Gly Gly  
Ala Arg Asp Thr Arg Pro Val Asp Pro Met Thr Gly Arg Tyr Val Asp  
Ala Glu Asn Lys Thr Val Asn Ala Asn Asp Asn Gln Pro Lys Glu Glu  
Asp Ser Glu Asp Lys Ile Ala  
120

35 (2) INFORMATION FOR SEQ ID NO:244:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 115 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..115  
(D) OTHER INFORMATION: / Ceres Seq. ID 1448135  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:244:  
aattctcag ggtttgaag agtgcctc gcgcggtg taattctct gtagcaaat  
cgacaaatg ggtcactcta atgtatggaa cttctatcgc aagaatcag gtcctggatc

40 (2) INFORMATION FOR SEQ ID NO:245:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 115 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..115  
(D) OTHER INFORMATION: / Ceres Seq. ID 1448015  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:245:  
Met Gly Leu Lys Phe Asp Arg Ile Lys Tyr Trp Leu Ser Val Gly Ala  
Gln Pro Ser Asp Pro Val Gln Arg Leu Leu Phe Arg Ser Gly Leu  
Pro Pro Pro Met Val Ala Met Gly Arg Lys Gly Gly Ala Arg Asp  
Thr Arg Pro Val Asp Pro Met Thr Gly Arg Tyr Val Asp Ala Glu Asn  
Lys Thr Val Asn Ala Asn Asp Asn Gln Pro Lys Glu Glu Asp Ser Glu  
Asp Lys Ile Ala  
70

45 (2) INFORMATION FOR SEQ ID NO:246:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 459 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..459  
(D) OTHER INFORMATION: / Ceres Seq. ID 1448135  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:246:  
aattctcag ggtttgaag agtgcctc gcgcggtg taattctct gtagcaaat  
cgacaaatg ggtcactcta atgtatggaa cttctatcgc aagaatcag gtcctggatc

50 (2) INFORMATION FOR SEQ ID NO:247:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 115 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..115  
(D) OTHER INFORMATION: / Ceres Seq. ID 1448135  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:247:  
aattctcag ggtttgaag agtgcctc gcgcggtg taattctct gtagcaaat  
cgacaaatg ggtcactcta atgtatggaa cttctatcgc aagaatcag gtcctggatc

55 (2) INFORMATION FOR SEQ ID NO:248:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 115 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..115  
(D) OTHER INFORMATION: / Ceres Seq. ID 1448135  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:248:  
aattctcag ggtttgaag agtgcctc gcgcggtg taattctct gtagcaaat  
cgacaaatg ggtcactcta atgtatggaa cttctatcgc aagaatcag gtcctggatc

60 (2) INFORMATION FOR SEQ ID NO:249:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 115 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..115  
(D) OTHER INFORMATION: / Ceres Seq. ID 1448135  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:249:  
aattctcag ggtttgaag agtgcctc gcgcggtg taattctct gtagcaaat  
cgacaaatg ggtcactcta atgtatggaa cttctatcgc aagaatcag gtcctggatc

439  
 ccgtttatgc cgtgtgtcgc ggaactcga cggctgtatc cgaagtatg qttgaactg 180  
 ccgaagacg tttcccgca gcaactgaa ggaattgga ttcattaaat accgttaate 240  
 aagaccacac ttaattgatt atgtttatg atataaacat gaagcgctgc aggggttgg 300  
 ctttaagct ttgg-agtt ttgaattt taatttgag aacctattt atttggag 360  
 ttaattaaat tttgaact cttactaac agctctatt tggattaat gattgtttgg 420  
 ctatttctg attttgatt taccatcaa attgact

(2) INFORMATION FOR SEQ ID NO:247:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 51 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ix) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..51

(D) OTHER INFORMATION: / Ceres Seq. ID 1448136

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:247:

Phe Ser Arg Val Lys Ser Val Ser Arg Arg Arg Cys Asn Ser Ser 10  
 1 Val Ser Lys Ser Thr Lys Trp Val Thr Leu Met Tyr Gly Thr Leu Ile 15  
 20 Arg Arg Ser Thr Val Leu Asp Leu Val Tyr Ala Val Cys Ala Gly Thr 30  
 35 Arg Thr Gly 40  
 50

(2) INFORMATION FOR SEQ ID NO:248:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 56 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ix) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..56

(D) OTHER INFORMATION: / Ceres Seq. ID 1448137

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:248:

Met Gly His Ser Asn Val Trp Asn Ser His Pro Lys Lys Tyr Gly Pro 10  
 1 Gly Ser Arg Leu Cys Arg Val Cys Gly Asn Ser His Gly Leu Ile Arg 15  
 20 Lys Tyr Gly Leu Asn Cys Cys Arg Gly Asn Phe Arg Ser Asn Ala -Lys 30  
 35 Glu Ile Gly Phe Ile Lys Tyr Arg 40  
 50

(2) INFORMATION FOR SEQ ID NO:249:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ix) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..38

(D) OTHER INFORMATION: / Ceres Seq. ID 1448138

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:249:

Met Glu Leu Ser Ser Glu Val Arg Ser Trp Ile Ser Phe Met Pro 10  
 1 Cys Val Arg Glu Leu Ala Arg Ala Asp Pro Glu Val Trp Phe Gly Leu 15  
 20

440

Leu Glu Thr Val Phe Pro 35

(2) INFORMATION FOR SEQ ID NO:250:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 603 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..603

(D) OTHER INFORMATION: / Ceres Seq. ID 1448185

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:250:

gaagaaga tatctgtg gagtttgg gatcgaaa atggagttc aggttcac 63  
 gaagaatg atcgacac aagaagat gctgcagt aaatagcac ttggtatag 120  
 agatgtgc gtcacatct tttccgct caacaaatc cgtcagctg agtttacct 180  
 tccatgaaa tgcagacy agcttcagt ttatgagag ttgatacag agcttttat 240  
 ggagaatg ttgcatac agaagatccg tgaagaaga gcttggcta aagaataa 300  
 actcaaga aacgtgctg ttctttat cctlaaact gctaa-gtt agattcgat 360  
 tctctcca attgcgat tccagattcc ggtattctct ggaac-gtga agatgggg 420  
 ggtctgct tcaatcttt ttctcttc ttctcttc ttctcttc ttctcttc 480  
 gatgaaatg tcaagattc tctctttt tttttttt gacctttt taacttggag 540  
 tggttccca aaaaaaaga tgaacaaact atctttttg tggtttata ttttaactc 600

(2) INFORMATION FOR SEQ ID NO:251:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 103 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ix) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..103

(D) OTHER INFORMATION: / Ceres Seq. ID 1448186

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:251:

Met Glu Val Pro Gly Ser Ser Lys Lys Met Ile Ala Thr Glu Glu 10  
 1 Leu Ser Ala Ala Lys Ile Ala Leu Gly Ser Arg Asp Met Cys Ala His 15  
 20 Leu Leu Ile Pro Leu Asn Lys Cys Arg Glu Ala Glu Phe Tyr Leu Pro 25  
 30 Trp Lys Cys Glu Asp Glu Arg His Val Tyr Glu Lys Cys Glu Tyr Glu 35  
 40 Leu Val Met Glu Arg Met Leu Ala Met Lys Lys Ile Arg Glu Glu 45  
 50 Ala Leu Ala Lys Glu Asn Lys Leu Glu Gly Asn Ala Ala Val Pro Leu 55  
 60 Ile Pro Lys Thr Ala Asn Ala 100

(2) INFORMATION FOR SEQ ID NO:252:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 94 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ix) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..94

(D) OTHER INFORMATION: / Ceres Seq. ID 1448187

441

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:252:

Met Ile Ala Thr Glu Glu Met Ser Ala Ala Lys Ile Ala Leu Gly 15  
 1 Ser Arg Asp Met Cys Ala His Leu Leu Ile Pro Leu Asn Lys Cys Arg 30  
 5 Gln Ala Glu Phe Tyr Leu Pro Trp Lys Cys Glu Asp Glu Arg His Val 45  
 35 Tyr Glu Lys Cys Glu Tyr Glu Leu Val Met Glu Arg Met Leu Ala Met 60  
 10 Lys Lys Ile Arg Glu Glu Ala Leu Ala Lys Lys Glu Asn Lys Leu Gln 75  
 65 Gly Asn Ala Ala Val Pro Ile Pro Lys Thr Ala Asn Ala 90

(2) INFORMATION FOR SEQ ID NO:253:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 87 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (D) LOCATION: 1..37

(D) OTHER INFORMATION: / Ceres Seq. ID 1448188

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:253:

Met Ser Ala Ala Lys Ile Ala Leu Gly Ser Arg Asp Met Cys Ala His 15  
 1 Leu Leu Ile Pro Leu Asn Lys Cys Arg Glu Gln Ala Glu Phe Tyr Leu Pro 30  
 20 Trp Lys Cys Glu Asp Glu Arg His Val Cys Cys Lys Cys Glu Tyr Glu 45  
 35 Leu Val Met Glu Arg Met Leu Ala Met Lys Lys Ile Arg Glu Glu Glu 60  
 65 Ala Leu Ala Lys Gln Asn Lys Leu Gln Gly Asn Ala Ala Val Pro Leu 80  
 10 Pro Lys Thr Ala Asn Ala 85

(2) INFORMATION FOR SEQ ID NO:254:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2034 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..2034

(D) OTHER INFORMATION: / Ceres Seq. ID 1450375

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:254:

ttatcaaa tgaatgctg ctcagatc ggttcgag ggcgagatg cgtgctctt 60  
 cactcaatg tcaatgtctc aaacgctgc ttagcgag tcttcgat tctggtcttc 120  
 tttctttct tctctctc aaacgctgc ttagcgatc cctgcgaatc aatctctac 180  
 aaactctag aagctcaag cagacccca tgaactaa ggaatgggt ttagctctg 240  
 ggtcttgag agagctgcta agctcttaq agatacaa agtctccc attcccaaa 300  
 tttctttat ctatagga agcagagaa aactcgtta gctgaattag cgcagagag 360  
 tttctttat gaagctatc agcagacaa ttaattggc agcagagaa aattgctga 420  
 gacccagaa aatctttgc agcagagc gcaaccaa gctcaaatc tgcataatg 480  
 ggaattgag gcaagaaag gacagagac gaatcatgaa gctcagagc atcaaatg 540  
 ggaattggt aagatgcaag aggtgtctc taccagaa gaaagagaa aaatgcacc 600  
 agaaagacg atccagctc agcagccta aactgagaa gaaagagctg aactgagc 660  
 agagagatt tptgtcagg cctagctga agctgaagt cgtgtcagc aagcaaatc 720  
 tactgaagc caaacagaa gattgcttat ggaagagatt aatggtgaa gaaagagtg 780

442

gcttgctga atcaaacaa tgttcagca catcgaagg ggaattcaga ccttataac 94C  
 tgaatgaat aagctgatta tgaatgttg aggcgtact gcaatgctc cagggttta 900  
 tacaactgt gaagagcta gattcatag ggttatatt ataggatgc ttggaacac 960  
 gtcaattt cgaagactt ccatggtag attctcatt agaggtcag tgtctaat 1020  
 5 taagaacagg attcaggag ctgacagc ttctcagca gaagcaaaa agcctctga 1080  
 taatgaatt ctcataact ctttgaaga acaatcagag cgtctccta gactcagc 1140  
 aaacaaaa tccataag caccattcg caactatg tttatagag cctctgtac 1200  
 cgttaaat atgttgacaa ggaatagc tccaatcg ggttgatt atcatatg 1260  
 gacagagt gatgtctc cctaggag acaatgtt acnaaatc atcagatatt 1320  
 10 tgaatggg agaatcga acnaaggtt accctttt atgatgag cgaatgtt 1380  
 tctatcag cgaacaga attatcag gaggctca cgaatgct tgaatgtt 1440  
 cctctcga atgttgatc aatctgga catgtttt gcttgctc caaacagac 1500  
 tggagatc: gatctggg tracaagag gattatga g-cattgat tccacttcc 1560  
 aggggaaga gaacttcc agctctcc tctctatc acaaatc taaagatgg 1620  
 tgataaac gaagacaa aaccgaatg ggtcaattg ttaagaag tgcacaga 1680  
 gattacgtt gaagagact taactgata agtgaattc gactgcaa agaagacaga 1740  
 aggtctct ggcg:aga ttgcaagct tctgctga gacaagct gactgacg 1800  
 acgagcagt cgtgttgg atcacagt tttaagag atgttaat ataatgta 1860  
 aqaacatc cgaagcata tctgtctc tgaagttt cagccatcc cctctctta 1920  
 gtcattga tttgattat acatgcat tactgtac gaagagaa tctgattgt 1980  
 ttaaaagag tctaaatg gaattttag attaaatg ttaagagt ttac

(2) INFORMATION FOR SEQ ID NO:255:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 639 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..639

(D) OTHER INFORMATION: / Ceres Seq. ID 1450876

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:255:

Leu Ser Lys Ser Met Ala Ala Ser Arg Leu Cys Ser Ala Ala Ala Ile 15  
 1 Ala Ala Ala Phe Thr Ser Met Ser Met Ser Gln Asn Arg Ala Tyr Ala 30  
 35 Asp Ser Ser Arg Phe Arg Phe Pro Phe Phe Ser Ser Ser Pro Ser Pro 45  
 40 Pro Pro Ser Asp Ser Pro Ala Asn Gln Ser Ser Thr Asn Ser Ser Lys 55  
 Ser Lys Ala Glu Pro Asp Glu Pro Lys Gly Ser Gly Phe Asp Pro Glu 65  
 65 Ala Leu Glu Arg Ala Ala Lys Ala Leu Arg Asp Ile Asn Ser Ser Pro 75  
 45 His Ser Lys Gln Val Phe Asp Leu Met Arg Lys Gln Glu Lys Thr Arg 85  
 100 Leu Ala Glu Leu Ala Ala Gln Thr Ser His Tyr Glu Ala Ile Gln Ala 110  
 115 His Asn Asp Ile Gly Arg Gln Gln Lys Leu Ala Glu Asp Gln Arg Asn 125  
 130 Leu Leu Gln Thr Gln Ala Gln Thr Lys Ala Gln Asn Leu Arg Tyr Glu 140  
 145 Asp Glu Leu Ala Arg Lys Arg Gln Gln Thr Asp His Glu Ala Arg 155  
 160 His His Asn Val Glu Leu Val Lys Met Gln Glu Ala Ser Ser Ile Arg 175  
 180 Lys Glu Lys Ala Lys Ile Ala Thr Glu Gln Ile Gln Ala Gln His 190  
 195 Arg Gln Thr Glu Lys Glu Arg Ala Glu Leu Gln Arg Glu Thr Ile Arg 205  
 210 Val Lys Ala Met Ala Glu Ala Glu Gly Arg Ala His Glu Ala Lys Leu 220  
 225



443

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:256:

Met Ala Ala Ser Arg Leu Cys Ser Ala Ala Ala Ile Ala Ala Ala Phe  
1 10 15  
Thr Ser Met Ser Met Ser Gln Asn Arg Ala Tyr Ala Asp Ser Ser Arg  
20 30  
Phe Arg Phe Phe Phe Ser Ser Ser Pro Ser Pro Pro Ser Asp  
35 40 45  
Ser Pro Ala Asn Gln Ser Ser Thr Asn Ser Ser Ser Ser Ser Ser Ser  
50 55 60  
Pro Asp Gln Pro Lys Gly Ser Gly Phe Asp Pro Gln Ala Leu Gln Arg  
65 70 75 80  
Ala Ala Lys Ala Leu Arg Asp Ile Asn Ser Ser Pro His Ser Lys Gln  
85 90  
Val Phe Asp Leu Met Arg Lys Gln Gln Lys Thr Arg Leu Ala Gln Leu  
100 105 110  
Ala Ala Gln Thr Ser His Tyr Gln Ala Ile Gln Ala His Asn Asp Ile  
115 120 125  
Gly Arg Gln Gln Lys Leu Ala Gln Asp Gln Arg Asn Leu Gln Thr  
130 135 140  
Gln Ala Gln Thr Lys Ala Gln Asn Leu Arg Tyr Gln Asp Gln Leu Ala  
145 150 155 160  
Arg Lys Arg Gln Gln Thr Asp His Gln Ala Gln Arg His His Asn Val  
165 170 175  
Glu Leu Val Lys Met Gln Gln Ala Ser Ser Ile Arg Lys Gln Lys Ala  
180 185 190  
Lys Ile Ala Thr Gln Gln Ile Gln Ala Gln His Arg Gln Thr Gln  
195 200 205  
Lys Gln Arg Ala Gln Leu Gln Arg Gln Thr Ile Arg Val Lys Ala Met  
210 215 220  
Ala Gln Ala Gln Gly Arg Ala His Gln Ala Lys Leu Thr Gln Gln Gln  
225 230 235 240  
Asn Arg Arg Leu Leu Met Gln Arg Ile Asn Gly Gln Arg Gln Lys Trp  
245 250 255  
Leu Ala Ala Ile Asn Thr Met Phe Ser His Ile Gln Gly Gly Phe Arg  
260 265 270  
Thr Leu Leu Thr Asp Arg Asn Lys Leu Ile Met Thr Val Gly Gly Ala  
275 280 285  
Thr Ala Leu Ala Ala Gly Val Tyr Thr Thr Arg Gln Gly Ala Arg Val  
290 295 300  
Thr Trp Gly Tyr Ile Asn Arg Met Leu Gly Gln Pro Ser Leu Ile Arg  
305 310 315 320  
Glu Ser Ser Met Arg Arg Phe Pro Trp Thr Gly Ser Val Ser Gln Phe  
325 330 335  
Lys Asn Arg Ile Ser Gly Ala Ala Ala Ser Ala Ala Gln Gly Lys  
340 345 350  
Lys Pro Leu Asp Asn Val Ile Leu His Thr Ser Leu Lys Arg Ile  
355 360 365  
Glu Arg Leu Ala Arg Ala Thr Ala Asn Thr Lys Ser His Gln Ala Pro  
370 375 380  
Phe Arg Asn Met Met Phe Tyr Gly Pro Pro Gly Thr Gly Lys Thr Met  
385 390 395  
Val Ala Arg Gln Ile Ala Arg Lys Ser Gly Leu Asp Tyr Ala Met Met  
400 405 410 415  
Thr Gly Gly Asp Val Ala Pro Leu Gly Ser Gln Ala Val Thr Lys Ile  
420 425 430  
His Gln Ile Phe Asp Trp Ala Lys Lys Ser Asn Arg Gly Leu Leu  
435 440 445  
Phe Ile Asp Gln Ala Asp Ala Phe Leu Cys Gln Arg Asn Ser Thr Tyr  
450 455 460  
Met Ser Gln Ala Gln Arg Ser Ala Leu Asn Ala Leu Phe Arg Thr  
465 470 475  
Gly Asn Gln Ser Arg Asp Ile Val Leu Val Leu Ala Thr Asn Arg Pro  
480 485 490 495 500

444

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:256:

Met Ala Ala Ser Arg Leu Cys Ser Ala Ala Ala Ile Ala Ala Ala Phe  
1 10 15  
Thr Ser Met Ser Met Ser Gln Asn Arg Ala Tyr Ala Asp Ser Ser Arg  
20 25 30  
Phe Arg Phe Phe Phe Ser Ser Ser Pro Ser Pro Pro Ser Asp  
35 40 45  
Ser Pro Ala Asn Gln Ser Ser Thr Asn Ser Ser Ser Ser Ser Ser Ser  
50 55 60  
Pro Asp Gln Pro Lys Gly Ser Gly Phe Asp Pro Gln Ala Leu Gln Arg  
65 70 75 80  
Ala Ala Lys Ala Leu Arg Asp Ile Asn Ser Ser Pro His Ser Lys Gln  
85 90  
Val Phe Asp Leu Met Arg Lys Gln Gln Lys Thr Arg Leu Ala Gln Leu  
100 105 110  
Ala Ala Gln Thr Ser His Tyr Gln Ala Ile Gln Ala His Asn Asp Ile  
115 120 125  
Gly Arg Gln Gln Lys Leu Ala Gln Asp Gln Arg Asn Leu Gln Thr  
130 135 140  
Gln Ala Gln Thr Lys Ala Gln Asn Leu Arg Tyr Gln Asp Gln Leu Ala  
145 150 155 160  
Arg Lys Arg Gln Gln Thr Asp His Gln Ala Gln Arg His His Asn Val  
165 170 175  
Glu Leu Val Lys Met Gln Gln Ala Ser Ser Ile Arg Lys Gln Lys Ala  
180 185 190  
Lys Ile Ala Thr Gln Gln Ile Gln Ala Gln His Arg Gln Thr Gln  
195 200 205  
Lys Gln Arg Ala Gln Leu Gln Arg Gln Thr Ile Arg Val Lys Ala Met  
210 215 220  
Ala Gln Ala Gln Gly Arg Ala His Gln Ala Lys Leu Thr Gln Gln Gln  
225 230 235 240  
Asn Arg Arg Leu Leu Met Gln Arg Ile Asn Gly Gln Arg Gln Lys Trp  
245 250 255  
Leu Ala Ala Ile Asn Thr Met Phe Ser His Ile Gln Gly Gly Phe Arg  
260 265 270  
Thr Leu Leu Thr Asp Arg Asn Lys Leu Ile Met Thr Val Gly Gly Ala  
275 280 285  
Thr Ala Leu Ala Ala Gly Val Tyr Thr Thr Arg Gln Gly Ala Arg Val  
290 295 300  
Thr Trp Gly Tyr Ile Asn Arg Met Leu Gly Gln Pro Ser Leu Ile Arg  
305 310 315 320  
Glu Ser Ser Met Arg Arg Phe Pro Trp Thr Gly Ser Val Ser Gln Phe  
325 330 335  
Lys Asn Arg Ile Ser Gly Ala Ala Ala Ser Ala Ala Gln Gly Lys  
340 345 350  
Lys Pro Leu Asp Asn Val Ile Leu His Thr Ser Leu Lys Arg Ile  
355 360 365  
Glu Arg Leu Ala Arg Ala Thr Ala Asn Thr Lys Ser His Gln Ala Pro  
370 375 380  
Phe Arg Asn Met Met Phe Tyr Gly Pro Pro Gly Thr Gly Lys Thr Met  
385 390 395  
Val Ala Arg Gln Ile Ala Arg Lys Ser Gly Leu Asp Tyr Ala Met Met  
400 405 410 415  
Thr Gly Gly Asp Val Ala Pro Leu Gly Ser Gln Ala Val Thr Lys Ile  
420 425 430  
His Gln Ile Phe Asp Trp Ala Lys Lys Ser Asn Arg Gly Leu Leu  
435 440 445  
Phe Ile Asp Gln Ala Asp Ala Phe Leu Cys Gln Arg Asn Ser Thr Tyr  
450 455 460  
Met Ser Gln Ala Gln Arg Ser Ala Leu Asn Ala Leu Phe Arg Thr  
465 470 475  
Gly Asn Gln Ser Arg Asp Ile Val Leu Val Leu Ala Thr Asn Arg Pro  
480 485 490 495 500

445

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:256:

Met Ala Ala Ser Arg Leu Cys Ser Ala Ala Ala Ile Ala Ala Ala Phe  
1 10 15  
Thr Ser Met Ser Met Ser Gln Asn Arg Ala Tyr Ala Asp Ser Ser Arg  
20 25 30  
Phe Arg Phe Phe Phe Ser Ser Ser Pro Ser Pro Pro Ser Asp  
35 40 45  
Ser Pro Ala Asn Gln Ser Ser Thr Asn Ser Ser Ser Ser Ser Ser Ser  
50 55 60  
Pro Asp Gln Pro Lys Gly Ser Gly Phe Asp Pro Gln Ala Leu Gln Arg  
65 70 75 80  
Ala Ala Lys Ala Leu Arg Asp Ile Asn Ser Ser Pro His Ser Lys Gln  
85 90  
Val Phe Asp Leu Met Arg Lys Gln Gln Lys Thr Arg Leu Ala Gln Leu  
100 105 110  
Ala Ala Gln Thr Ser His Tyr Gln Ala Ile Gln Ala His Asn Asp Ile  
115 120 125  
Gly Arg Gln Gln Lys Leu Ala Gln Asp Gln Arg Asn Leu Gln Thr  
130 135 140  
Gln Ala Gln Thr Lys Ala Gln Asn Leu Arg Tyr Gln Asp Gln Leu Ala  
145 150 155 160  
Arg Lys Arg Gln Gln Thr Asp His Gln Ala Gln Arg His His Asn Val  
165 170 175  
Glu Leu Val Lys Met Gln Gln Ala Ser Ser Ile Arg Lys Gln Lys Ala  
180 185 190  
Lys Ile Ala Thr Gln Gln Ile Gln Ala Gln His Arg Gln Thr Gln  
195 200 205  
Lys Gln Arg Ala Gln Leu Gln Arg Gln Thr Ile Arg Val Lys Ala Met  
210 215 220  
Ala Gln Ala Gln Gly Arg Ala His Gln Ala Lys Leu Thr Gln Gln Gln  
225 230 235 240  
Asn Arg Arg Leu Leu Met Gln Arg Ile Asn Gly Gln Arg Gln Lys Trp  
245 250 255  
Leu Ala Ala Ile Asn Thr Met Phe Ser His Ile Gln Gly Gly Phe Arg  
260 265 270  
Thr Leu Leu Thr Asp Arg Asn Lys Leu Ile Met Thr Val Gly Gly Ala  
275 280 285  
Thr Ala Leu Ala Ala Gly Val Tyr Thr Thr Arg Gln Gly Ala Arg Val  
290 295 300  
Thr Trp Gly Tyr Ile Asn Arg Met Leu Gly Gln Pro Ser Leu Ile Arg  
305 310 315 320  
Glu Ser Ser Met Arg Arg Phe Pro Trp Thr Gly Ser Val Ser Gln Phe  
325 330 335  
Lys Asn Arg Ile Ser Gly Ala Ala Ala Ser Ala Ala Gln Gly Lys  
340 345 350  
Lys Pro Leu Asp Asn Val Ile Leu His Thr Ser Leu Lys Arg Ile  
355 360 365  
Glu Arg Leu Ala Arg Ala Thr Ala Asn Thr Lys Ser His Gln Ala Pro  
370 375 380  
Phe Arg Asn Met Met Phe Tyr Gly Pro Pro Gly Thr Gly Lys Thr Met  
385 390 395  
Val Ala Arg Gln Ile Ala Arg Lys Ser Gly Leu Asp Tyr Ala Met Met  
400 405 410 415  
Thr Gly Gly Asp Val Ala Pro Leu Gly Ser Gln Ala Val Thr Lys Ile  
420 425 430  
His Gln Ile Phe Asp Trp Ala Lys Lys Ser Asn Arg Gly Leu Leu  
435 440 445  
Phe Ile Asp Gln Ala Asp Ala Phe Leu Cys Gln Arg Asn Ser Thr Tyr  
450 455 460  
Met Ser Gln Ala Gln Arg Ser Ala Leu Asn Ala Leu Phe Arg Thr  
465 470 475  
Gly Asn Gln Ser Arg Asp Ile Val Leu Val Leu Ala Thr Asn Arg Pro  
480 485 490 495 500

446

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:256:

Met Ala Ala Ser Arg Leu Cys Ser Ala Ala Ala Ile Ala Ala Ala Phe  
1 10 15  
Thr Ser Met Ser Met Ser Gln Asn Arg Ala Tyr Ala Asp Ser Ser Arg  
20 25 30  
Phe Arg Phe Phe Phe Ser Ser Ser Pro Ser Pro Pro Ser Asp  
35 40 45  
Ser Pro Ala Asn Gln Ser Ser Thr Asn Ser Ser Ser Ser Ser Ser Ser  
50 55 60  
Pro Asp Gln Pro Lys Gly Ser Gly Phe Asp Pro Gln Ala Leu Gln Arg  
65 70 75 80  
Ala Ala Lys Ala Leu Arg Asp Ile Asn Ser Ser Pro His Ser Lys Gln  
85 90  
Val Phe Asp Leu Met Arg Lys Gln Gln Lys Thr Arg Leu Ala Gln Leu  
100 105 110  
Ala Ala Gln Thr Ser His Tyr Gln Ala Ile Gln Ala His Asn Asp Ile  
115 120 125  
Gly Arg Gln Gln Lys Leu Ala Gln Asp Gln Arg Asn Leu Gln Thr  
130 135 140  
Gln Ala Gln Thr Lys Ala Gln Asn Leu Arg Tyr Gln Asp Gln Leu Ala  
145 150 155 160  
Arg Lys Arg Gln Gln Thr Asp His Gln Ala Gln Arg His His Asn Val  
165 170 175  
Glu Leu Val Lys Met Gln Gln Ala Ser Ser Ile Arg Lys Gln Lys Ala  
180 185 190  
Lys Ile Ala Thr Gln Gln Ile Gln Ala Gln His Arg Gln Thr Gln  
195 200 205  
Lys Gln Arg Ala Gln Leu Gln Arg Gln Thr Ile Arg Val Lys Ala Met  
210 215 220  
Ala Gln Ala Gln Gly Arg Ala His Gln Ala Lys Leu Thr Gln Gln Gln  
225 230 235 24

445 490 495  
Gly Asp Leu Asp Ser Ala Val Thr Asp Arg Ile Asp Glu Val Ile Glu 500  
Phe Pro Leu Pro Gly Glu Glu Arg Phe Lys Leu Leu Asn Leu Tyr 515  
Leu Asn Lys Tyr Leu Lys Met Gly Asp Asn Asn Glu Asp Thr Lys Pro 530  
Lys Trp Ser His Leu Phe Lys Lys Ieu Ser Gln Lys Ile Thr Val Glu 545  
Glu Asp Leu Thr Asp Lys Val Ile Ser Glu Ala Ala Lys Lys Thr Glu 560  
Gly Phe Ser Gly Arg Glu Ile Ala Lys Leu Val Ala Gly Val Gln Ala 580  
Gly Val Tyr Gly Arg Ala Asp Cys Val Leu Asp Ser Gln Leu Phe Lys 595  
Glu Ile Val Glu Tyr Lys Val Glu Glu His His Arg Arg His Met Leu 610  
Ala Ser Gln Gly Phe Gln Pro Leu Leu Phe Ser 620  
625 630 635  
(2) INFORMATION FOR SEQ ID NO:257:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 617 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..617  
(D) OTHER INFORMATION: / Ceres Seq. ID 1450878  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:257:  
Met Ser Met Ser Gln Asn Arg Ala Tyr Ala Asp Ser Ser Arg Phe Arg 1  
Phe Pro Phe Ser Ser Ser Pro Ser Pro Pro Ser Asp Ser Pro 5  
Ala Asn Gln Ser Ser Thr Asn Ser Ser Lys Ser Lys Ala Glu Pro Asp 20  
Glu Pro Lys Gly Ser Gly Phe Asp Pro Glu Ala Leu Cys Arg Ala Ala 35  
Lys Ala Leu Arg Asp Ile Asn Ser Ser Pro His Ser Lys Gln Val Phe 50  
Asp Leu Met Arg Lys Gln Glu Lys Thr Arg Leu Ala Glu Leu Ala Ala 65  
Glu Thr Ser His Tyr Glu Ala Ile Gln Ala His Asn Asp Ile Gly Arg 80  
Gln Gln Lys Leu Ala Glu Asp Gln Arg Asn Leu Leu Gln Thr Gln Ala 95  
Gln Thr Lys Ala Gln Asn Leu Arg Tyr Glu Asp Glu Leu Ala Arg Lys 110  
Arg Gln Gln Thr Asp His Glu Ala Gln Arg His Asn Val Glu Leu 125  
Val Lys Met Gln Glu Ala Ser Ser Ile Arg Lys Glu Lys Ala Lys Ile 140  
Ala Thr Glu Glu Gln Ile Gln Ala His Arg Gln Thr Glu Lys Glu 155  
Arg Ala Glu Leu Glu Arg Glu Thr Ile Arg Val Lys Ala Met Ala Glu 170  
Ala Glu Gly Arg Ala His Glu Ala Lys Ieu Thr Glu Gln Asn Arg 185  
Arg Leu Leu Met Glu Arg Ile Asn Gly Glu Arg Glu Lys Trp Leu Ala 200  
Ala Ile Asn Thr Met Phe Ser His Ile Glu Gly Gly Phe Arg Thr Leu 215  
220 225 230 235 240

446 250 255  
Leu Thr Asp Arg Asn Lys Leu Ile Met Thr Val Gly Gly Ala Thr Ala 260  
Leu Ala Ala Gly Val Tyr Thr Thr Arg Glu Gly Ala Arg Val Thr Trp 270  
Gly Tyr Ile Asn Arg Met Leu Gly Gln Pro Ser Leu Ile Arg Glu Ser 285  
Ser Met Arg Arg Phe Pro Trp Thr Gly Ser Val Ser Gln Phe Lys Asn 300  
Arg Ile Ser Gly Ala Ala Ala Ser Ala Ala Glu Gly Lys Lys Pro 315  
Leu Asp Asn Val Ile Leu His Thr Ser Leu Lys Lys Arg Ile Glu Arg 330  
Leu Ala Arg Ala Thr Ala Asn Thr Lys Ser His Gln Ala Pro Phe Arg 345  
Asn Met Met Phe Tyr Gly Pro Pro Gly Thr Gly Lys Thr Met Val Ala 360  
Arg Glu Ile Ala Arg Lys Ser Gly Leu Asp Tyr Ala Met Met Thr Gly 375  
Gly Asp Val Ala Pro Leu Gly Ser Gln Ala Val Thr Lys Ile His Gln 390  
Ile Phe Asp Trp Ala Lys Lys Ser Asn Arg Gly Leu Leu Phe Ile 405  
Asp Glu Ala Asp Ala Phe Leu Cys Glu Arg Asn Ser Thr Tyr Met Ser 420  
Glu Ala Gln Arg Ser Ala Leu Asn Ala Leu Phe Arg Thr Gly Asp 435  
Gln Ser Arg Asp Ile Val Leu Val Leu Ala Thr Asn Arg Pro Gly Asp 450  
Leu Asp Ser Ala Val Thr Asp Arg Ile Asp Glu Val Ile Glu Phe Pro 465  
Leu Pro Gly Glu Glu Arg Phe Lys Leu Leu Leu Asn Leu Tyr Leu Asn 480  
Lys Tyr Leu Lys Met Gly Asp Asn Asn Glu Asp Thr Lys Pro Lys Trp 495  
Ser His Leu Phe Lys Lys Leu Ser Gln Lys Ile Thr Val Glu Glu Asp 510  
Leu Thr Asp Lys Val Ile Ser Glu Ala Ala Lys Lys Thr Glu Gly Phe 525  
Ser Gly Arg Glu Ile Ala Lys Leu Val Ala Gly Val Gln Ala Gly Val 540  
Tyr Gly Arg Ala Asp Cys Val Leu Asp Ser Gln Leu Phe Lys Glu Ile 555  
Val Glu Tyr Lys Val Glu Glu His His Arg Arg His Met Leu Ala Ser 570  
Glu Gly Phe Gln Pro Leu Phe Ser 585  
600 605  
(2) INFORMATION FOR SEQ ID NO:258:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 478 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..478  
(D) OTHER INFORMATION: / Ceres Seq. ID 1459191  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:258:  
aacctcttg gaaagtgct caactctgc agagaaaag aacaagaag atcccgaaa 60  
atggcaagg cgatgtacg tccagcagc agctccgagc agtccgaga 120  
catcctgcg tctaagcga aactttctc ctccgcgg ccagcgat gcttatgaag 180

447  
ctgcgaagtg ggaagaata actatctctg gttctctg ttcactctg ctactctg  
atgtttatc caagggcat catctggcg gagaaga gttctcttg gttcttgatg  
gtcgttga gttgaagac acaaaagc actgagctt gctggtcat ataaagct  
tcagggtta ttgaagagc taaatgtt taccgtattt gttccacg ttgtcaacg  
5 attactact ccaactctt tctcttgtt ggaataaa agttaact ttgtctg  
(2) INFORMATION FOR SEQ ID NO:259:  
(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 66 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS:  
    (ii) MOLECULE TYPE: linear  
    (iii) MOLECULE TYPE: peptide  
    (iv) FEATURE:  
        (A) NAME/KEY: peptide  
        (B) LOCATION: 1..66  
(D) OTHER INFORMATION: / Ceres Seq. ID 1459192  
Asn Pro Leu Gly Lys Ser Leu Asn Thr Lys Arg Glu Lys Glu Gln Gly  
1  
20 Arg Ser Arg Lys Met Ala Thr Ala Ile Val Arg Ser Ala Leu Ser Arg  
20  
Ala Val Thr Arg Ala Val Arg Arg His Pro Ser Leu Leu Ser Glu Thr  
35  
Phe Pro Leu Pro Ala Met Thr Met Leu Met Lys Leu Arg Ser Gly  
50  
Arg Arg  
65  
(2) INFORMATION FOR SEQ ID NO:260:  
(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 110 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS:  
    (ii) MOLECULE TYPE: linear  
    (iii) MOLECULE TYPE: peptide  
    (iv) FEATURE:  
        (A) NAME/KEY: peptide  
        (B) LOCATION: 1..110  
(D) OTHER INFORMATION: / Ceres Seq. ID 1459193  
Pro Ser Trp Lys Glu Ser Gln His Leu Gln Arg Lys Arg Thr Arg Lys  
1  
Ile Pro Glu Asn Gly Asn Gly Asp Cys Thr Phe Ser Ser Phe Pro Ser  
20  
Ser Asp Ser Arg Ser Pro Lys Thr Ser Val Ala Pro Lys Arg Asn Phe  
35  
Ser Ser Ala Gly His Asp Asp Ala Tyr Glu Ala Ala Lys Trp Glu  
50  
Lys Ile Thr Tyr Leu Gly Ile Ala Ser Cys Thr Ala Leu Ala Val Tyr  
65  
Val Leu Ser Lys Gly His His His Gly Glu Asp Lys Glu Phe Pro Trp  
85  
Gly Pro Asp Gly Leu Phe Glu Val Lys His Asn Lys Glu His  
110  
(2) INFORMATION FOR SEQ ID NO:261:  
(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 741 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (ii) MOLECULE TYPE: DNA (genomic)  
    (iii) FEATURE:  
        (A) NAME/KEY: -

448  
(B) LOCATION: 1..741  
(D) OTHER INFORMATION: / Ceres Seq. ID 1461848  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:261:  
attttagtac attgtgacc attcttttg tatagactac tatctctgat ctcttgagag  
60  
t-aagtcac aactaggaaa attcagaagc gcttcaaac tcaaaaalal ccatggcgag  
120  
gattacagaa ttctaccaa aagagtcacq atagtcgrt ctgcctctcg tctctactg  
180  
ttctctcac ctctgagcg gtcgtcaagt cggcagagct cgaagaaggt acaacgtccc  
240  
gtatccaact ctatgtcaa tagaatcaga aacaagaat gctaaagctc tcaactgtgt  
300  
tcaagagga catcaaaact cttagagat gatccaatg tatttcatac tgatgacct  
360  
cggggagtg aagcaacatt gtatcttac tggccttgt tgccttaca acgttagccq  
420  
attctctac tttaaggtt atgcctcttg agatccatg atctcttg gttgacct  
480  
atcgggttc ttgggttgc taggtctgat gatatgacc atctcttg gttgacct  
540  
gatacttgt taagctactc gttctgggg ttaagatgc tctggttgc tgaagaata  
600  
tgaacacat gctgtgaagc tgcacaaa acttggttaa tacttagag ttctgacct  
660  
tcaaaactt ttaataatc atggtctcat agaacagttg aaatttcac tccgtagacg  
720  
tcaataaga ttgaattat g  
(2) INFORMATION FOR SEQ ID NO:262:  
(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 146 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS:  
    (ii) MOLECULE TYPE: linear  
    (iii) MOLECULE TYPE: peptide  
    (iv) FEATURE:  
        (A) NAME/KEY: peptide  
        (B) LOCATION: 1..146  
(D) OTHER INFORMATION: / Ceres Seq. ID 1461849  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:262:  
Met Ala Ala Ile Thr Glu Phe Leu Pro Lys Glu Tyr Gly Tyr Val Val  
1  
Leu Val Leu Val Phe Tyr Cys Phe Leu Asn Leu Trp Met Gly Ala Gln  
5  
Val Gly Arg Ala Arg Lys Arg Tyr Asn Val Pro Tyr Pro Thr Leu Tyr  
10  
Ala Ile Glu Ser Glu Asn Lys Asp Ala Lys Leu Phe Asn Cys Val Gln  
15  
Arg Gly His Gln Asn Ser Leu Glu Met Met Pro Met Tyr Phe Ile Leu  
20  
Met Ile Leu Gly Gly Met Lys His Pro Cys Ile Cys Thr Gly Leu Gly  
25  
Leu Leu Tyr Asn Val Ser Arg Phe Phe Tyr Phe Lys Gly Tyr Ala Thr  
30  
Gly Asp Pro Met Lys Arg Leu Thr Ile Gly Lys Tyr Gly Phe Leu Gly  
35  
Leu Leu Gly Leu Met Ile Cys Thr Ile Ser Phe Gly Val Thr Ile Ile  
40  
Leu Ala  
45  
Leu Ala  
50  
(2) INFORMATION FOR SEQ ID NO:263:  
(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 118 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS:  
    (ii) MOLECULE TYPE: linear  
    (iii) MOLECULE TYPE: peptide  
    (iv) FEATURE:  
        (A) NAME/KEY: peptide  
        (B) LOCATION: 1..118  
(D) OTHER INFORMATION: / Ceres Seq. ID 1461850  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:263:  
Met Gly Ala Gln Val Gly Arg Ala Arg Lys Arg Tyr Asn Val Pro Tyr  
1  
10  
15

449  
 Pro Thr Leu Tyr Ala Ile Glu Ser Glu Asn Lys Asp Ala Lys Leu Phe 30  
 Asn Cys Val Gln Arg Gly His Gln Asn Ser Leu Glu Met Met Pro Met 45  
 5 Tyr Phe Ile Leu Met Ile Leu Gly Gly Met Lys His Pro Cys Ile Cys 50  
 Thr Gly Leu Gly Leu Tyr Asn Val Ser Arg Phe Tyr Phe Lys 60  
 65 Gly Tyr Ala Thr Gly Asp Pro Met Lys Arg Leu Thr Ile Gly Lys Tyr 75  
 10 Gly Phe Leu Gly Leu Gly Leu Met Ile Cys Thr Ile Ser Phe Gly 80  
 100 Val Thr Leu Ile Leu Ala 105  
 115

15 (2) INFORMATION FOR SEQ ID NO:264:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 74 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..74  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1461851

25 Met Met Pro Met Tyr Phe Ile Leu Met Ile Leu Gly Gly Met Lys His 1  
 5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55  
 60  
 65  
 70  
 75  
 80  
 85  
 90  
 95  
 100  
 105  
 110  
 115  
 120  
 125  
 130  
 135  
 140  
 145  
 150  
 155  
 160  
 165  
 170  
 175  
 180  
 185  
 190  
 195  
 200  
 205  
 210  
 215  
 220  
 225  
 230  
 235  
 240  
 245  
 250  
 255  
 260  
 265  
 270  
 275  
 280  
 285  
 290  
 295  
 300  
 305  
 310  
 315  
 320  
 325  
 330  
 335  
 340  
 345  
 350  
 355  
 360  
 365  
 370  
 375  
 380  
 385  
 390  
 395  
 400  
 405  
 410  
 415  
 420  
 425  
 430  
 435  
 440  
 445  
 450  
 455  
 460  
 465  
 470  
 475  
 480  
 485  
 490  
 495  
 500  
 505  
 510  
 515  
 520  
 525  
 530  
 535  
 540  
 545  
 550  
 555  
 560  
 565  
 570  
 575  
 580  
 585  
 590  
 595  
 600  
 605  
 610  
 615  
 620  
 625  
 630  
 635  
 640  
 645  
 650  
 655  
 660  
 665  
 670  
 675  
 680  
 685  
 690  
 695  
 700  
 705  
 710  
 715  
 720  
 725  
 730  
 735  
 740  
 745  
 750  
 755  
 760  
 765  
 770  
 775  
 780  
 785  
 790  
 795  
 800  
 805  
 810  
 815  
 820  
 825  
 830  
 835  
 840  
 845  
 850  
 855  
 860  
 865  
 870  
 875  
 880  
 885  
 890  
 895  
 900  
 905  
 910  
 915  
 920  
 925  
 930  
 935  
 940  
 945  
 950  
 955  
 960  
 965  
 970  
 975  
 980  
 985  
 990  
 995  
 1000  
 1005  
 1010  
 1015  
 1020  
 1025  
 1030  
 1035  
 1040  
 1045  
 1050  
 1055  
 1060  
 1065  
 1070  
 1075  
 1080  
 1085  
 1090  
 1095  
 1100  
 1105  
 1110  
 1115  
 1120  
 1125  
 1130  
 1135  
 1140  
 1145  
 1150  
 1155  
 1160  
 1165  
 1170  
 1175  
 1180  
 1185  
 1190  
 1195  
 1200  
 1205  
 1210  
 1215  
 1220  
 1225  
 1230  
 1235  
 1240  
 1245  
 1250  
 1255  
 1260  
 1265  
 1270  
 1275  
 1280  
 1285  
 1290  
 1295  
 1300  
 1305  
 1310  
 1315  
 1320  
 1325  
 1330  
 1335  
 1340  
 1345  
 1350  
 1355  
 1360  
 1365  
 1370  
 1375  
 1380  
 1385  
 1390  
 1395  
 1400  
 1405  
 1410  
 1415  
 1420  
 1425  
 1430  
 1435  
 1440  
 1445  
 1450  
 1455  
 1460  
 1465  
 1470  
 1475  
 1480  
 1485  
 1490  
 1495  
 1500  
 1505  
 1510  
 1515  
 1520  
 1525  
 1530  
 1535  
 1540  
 1545  
 1550  
 1555  
 1560  
 1565  
 1570  
 1575  
 1580  
 1585  
 1590  
 1595  
 1600  
 1605  
 1610  
 1615  
 1620  
 1625  
 1630  
 1635  
 1640  
 1645  
 1650  
 1655  
 1660  
 1665  
 1670  
 1675  
 1680  
 1685  
 1690  
 1695  
 1700  
 1705  
 1710  
 1715  
 1720  
 1725  
 1730  
 1735  
 1740  
 1745  
 1750  
 1755  
 1760  
 1765  
 1770  
 1775  
 1780  
 1785  
 1790  
 1795  
 1800  
 1805  
 1810  
 1815  
 1820  
 1825  
 1830  
 1835  
 1840  
 1845  
 1850  
 1855  
 1860  
 1865  
 1870  
 1875  
 1880  
 1885  
 1890  
 1895  
 1900  
 1905  
 1910  
 1915  
 1920  
 1925  
 1930  
 1935  
 1940  
 1945  
 1950  
 1955  
 1960  
 1965  
 1970  
 1975  
 1980  
 1985  
 1990  
 1995  
 2000  
 2005  
 2010  
 2015  
 2020  
 2025  
 2030  
 2035  
 2040  
 2045  
 2050  
 2055  
 2060  
 2065  
 2070  
 2075  
 2080  
 2085  
 2090  
 2095  
 2100  
 2105  
 2110  
 2115  
 2120  
 2125  
 2130  
 2135  
 2140  
 2145  
 2150  
 2155  
 2160  
 2165  
 2170  
 2175  
 2180  
 2185  
 2190  
 2195  
 2200  
 2205  
 2210  
 2215  
 2220  
 2225  
 2230  
 2235  
 2240  
 2245  
 2250  
 2255  
 2260  
 2265  
 2270  
 2275  
 2280  
 2285  
 2290  
 2295  
 2300  
 2305  
 2310  
 2315  
 2320  
 2325  
 2330  
 2335  
 2340  
 2345  
 2350  
 2355  
 2360  
 2365  
 2370  
 2375  
 2380  
 2385  
 2390  
 2395  
 2400  
 2405  
 2410  
 2415  
 2420  
 2425  
 2430  
 2435  
 2440  
 2445  
 2450  
 2455  
 2460  
 2465  
 2470  
 2475  
 2480  
 2485  
 2490  
 2495  
 2500  
 2505  
 2510  
 2515  
 2520  
 2525  
 2530  
 2535  
 2540  
 2545  
 2550  
 2555  
 2560  
 2565  
 2570  
 2575  
 2580  
 2585  
 2590  
 2595  
 2600  
 2605  
 2610  
 2615  
 2620  
 2625  
 2630  
 2635  
 2640  
 2645  
 2650  
 2655  
 2660  
 2665  
 2670  
 2675  
 2680  
 2685  
 2690  
 2695  
 2700  
 2705  
 2710  
 2715  
 2720  
 2725  
 2730  
 2735  
 2740  
 2745  
 2750  
 2755  
 2760  
 2765  
 2770  
 2775  
 2780  
 2785  
 2790  
 2795  
 2800  
 2805  
 2810  
 2815  
 2820  
 2825  
 2830  
 2835  
 2840  
 2845  
 2850  
 2855  
 2860  
 2865  
 2870  
 2875  
 2880  
 2885  
 2890  
 2895  
 2900  
 2905  
 2910  
 2915  
 2920  
 2925  
 2930  
 2935  
 2940  
 2945  
 2950  
 2955  
 2960  
 2965  
 2970  
 2975  
 2980  
 2985  
 2990  
 2995  
 3000  
 3005  
 3010  
 3015  
 3020  
 3025  
 3030  
 3035  
 3040  
 3045  
 3050  
 3055  
 3060  
 3065  
 3070  
 3075  
 3080  
 3085  
 3090  
 3095  
 3100  
 3105  
 3110  
 3115  
 3120  
 3125  
 3130  
 3135  
 3140  
 3145  
 3150  
 3155  
 3160  
 3165  
 3170  
 3175  
 3180  
 3185  
 3190  
 3195  
 3200  
 3205  
 3210  
 3215  
 3220  
 3225  
 3230  
 3235  
 3240  
 3245  
 3250  
 3255  
 3260  
 3265  
 3270  
 3275  
 3280  
 3285  
 3290  
 3295  
 3300  
 3305  
 3310  
 3315  
 3320  
 3325  
 3330  
 3335  
 3340  
 3345  
 3350  
 3355  
 3360  
 3365  
 3370  
 3375  
 3380  
 3385  
 3390  
 3395  
 3400  
 3405  
 3410  
 3415  
 3420  
 3425  
 3430  
 3435  
 3440  
 3445  
 3450  
 3455  
 3460  
 3465  
 3470  
 3475  
 3480  
 3485  
 3490  
 3495  
 3500  
 3505  
 3510  
 3515  
 3520  
 3525  
 3530  
 3535  
 3540  
 3545  
 3550  
 3555  
 3560  
 3565  
 3570  
 3575  
 3580  
 3585  
 3590  
 3595  
 3600  
 3605  
 3610  
 3615  
 3620  
 3625  
 3630  
 3635  
 3640  
 3645  
 3650  
 3655  
 3660  
 3665  
 3670  
 3675  
 3680  
 3685  
 3690  
 3695  
 3700  
 3705  
 3710  
 3715  
 3720  
 3725  
 3730  
 3735  
 3740  
 3745  
 3750  
 3755  
 3760  
 3765  
 3770  
 3775  
 3780  
 3785  
 3790  
 3795  
 3800  
 3805  
 3810  
 3815  
 3820  
 3825  
 3830  
 3835  
 3840  
 3845  
 3850  
 3855  
 3860  
 3865  
 3870  
 3875  
 3880  
 3885  
 3890  
 3895  
 3900  
 3905  
 3910  
 3915  
 3920  
 3925  
 3930  
 3935  
 3940  
 3945  
 3950  
 3955  
 3960  
 3965  
 3970  
 3975  
 3980  
 3985  
 3990  
 3995  
 4000  
 4005  
 4010  
 4015  
 4020  
 4025  
 4030  
 4035  
 4040  
 4045  
 4050  
 4055  
 4060  
 4065  
 4070  
 4075  
 4080  
 4085  
 4090  
 4095  
 4100  
 4105  
 4110  
 4115  
 4120  
 4125  
 4130  
 4135  
 4140  
 4145  
 4150  
 4155  
 4160  
 4165  
 4170  
 4175  
 4180  
 4185  
 4190  
 4195  
 4200  
 4205  
 4210  
 4215  
 4220  
 4225  
 4230  
 4235  
 4240  
 4245  
 4250  
 4255  
 4260  
 4265  
 4270  
 4275  
 4280  
 4285  
 4290  
 4295  
 4300  
 4305  
 4310  
 4315  
 4320  
 4325  
 4330  
 4335  
 4340  
 4345  
 4350  
 4355  
 4360  
 4365  
 4370  
 4375  
 4380  
 4385  
 4390  
 4395  
 4400  
 4405  
 4410  
 4415  
 4420  
 4425  
 4430  
 4435  
 4440  
 4445  
 4450  
 4455  
 4460  
 4465  
 4470  
 4475  
 4480  
 4485  
 4490  
 4495  
 4500  
 4505  
 4510  
 4515  
 4520  
 4525  
 4530  
 4535  
 4540  
 4545  
 4550  
 4555  
 4560  
 4565  
 4570  
 4575  
 4580  
 4585  
 4590  
 4595  
 4600  
 4605  
 4610  
 4615  
 4620  
 4625  
 4630  
 4635  
 4640  
 4645  
 4650  
 4655  
 4660  
 4665  
 4670  
 4675  
 4680  
 4685  
 4690  
 4695  
 4700  
 4705  
 4710  
 4715  
 4720  
 4725  
 4730  
 4735  
 4740  
 4745  
 4750  
 4755  
 4760  
 4765  
 4770  
 4775  
 4780  
 4785  
 4790  
 4795  
 4800  
 4805  
 4810  
 4815  
 4820  
 4825  
 4830  
 4835  
 4840  
 4845  
 4850  
 4855  
 4860  
 4865  
 4870  
 4875  
 4880  
 4885  
 4890  
 4895  
 4900  
 4905  
 4910  
 4915  
 4920  
 4925  
 4930  
 4935  
 4940  
 4945  
 4950  
 4955  
 4960  
 4965  
 4970  
 4975  
 4980  
 4985  
 4990  
 4995  
 5000  
 5005  
 5010  
 5015  
 5020  
 5025  
 5030  
 5035  
 5040  
 5045  
 5050  
 5055  
 5060  
 5065  
 5070  
 5075  
 5080  
 5085  
 5090  
 5095  
 5100  
 5105  
 5110  
 5115  
 5120  
 5125  
 5130  
 5135  
 5140  
 5145  
 5150  
 5155  
 5160  
 5165  
 5170  
 5175  
 5180  
 5185  
 5190  
 5195  
 5200  
 5205  
 5210  
 5215  
 5220  
 5225  
 5230  
 5235  
 5240  
 5245  
 5250  
 5255  
 5260  
 5265  
 5270  
 5275  
 5280  
 5285  
 5290  
 5295  
 5300  
 5305  
 5310  
 5315  
 5320  
 5325  
 5330  
 5335  
 5340  
 5345  
 5350  
 5355  
 5360  
 5365  
 5370  
 5375  
 5380  
 5385  
 5390  
 5395  
 5400  
 5405  
 5410  
 5415  
 5420  
 5425  
 5430  
 5435  
 5440  
 5445  
 5450  
 5455  
 5460  
 5465  
 5470  
 5475  
 5480  
 5485  
 5490  
 5495  
 5500  
 5505  
 5510  
 5515  
 5520  
 5525  
 5530  
 5535  
 5540  
 5545  
 5550  
 5555  
 5560  
 5565  
 5570  
 5575  
 5580  
 5585  
 5590  
 5595  
 5600  
 5605  
 5610  
 5615  
 5620  
 5625  
 5630  
 5635  
 5640  
 5645  
 5650  
 5655  
 5660  
 5665  
 5670  
 5675  
 5680  
 5685  
 5690  
 5695  
 5700  
 5705  
 5710  
 5715  
 5720  
 5725  
 5730  
 5735  
 5740  
 5745  
 5750  
 5755  
 5760  
 5765  
 5770  
 5775  
 5780  
 5785  
 5790  
 5795  
 5800  
 5805  
 5810  
 5815  
 5820  
 5825  
 5830  
 5835  
 5840  
 5845  
 5850  
 5855  
 5860  
 5865  
 5870  
 5875  
 5880  
 5885  
 5890  
 5895  
 5900  
 5905  
 5910  
 5915  
 5920  
 5925  
 5930  
 5935  
 5940  
 5945  
 5950  
 5955  
 5960  
 5965  
 5970  
 5975  
 5980  
 5985  
 5990  
 5995  
 6000  
 6005  
 6010  
 6015  
 6020  
 6025  
 6030  
 6035  
 6040  
 6045  
 6050  
 6055  
 6060  
 6065  
 6070  
 6075  
 6080  
 6085  
 6090  
 6095  
 6100  
 6105  
 6110  
 6115  
 6120  
 6125  
 6130  
 6135  
 6140  
 6145  
 6150  
 6155  
 6160  
 6165  
 6170  
 6175  
 6180  
 6185  
 6190  
 6195  
 6200  
 6205  
 6210  
 6215  
 6220  
 6225  
 6230  
 6235  
 6240  
 6245  
 6250  
 6255  
 6260  
 6265  
 6270  
 6275  
 6280  
 6285  
 6290  
 6295  
 6300  
 6305  
 6310  
 6315  
 6320  
 6325  
 6330  
 6335  
 6340  
 6345  
 6350  
 6355  
 6360  
 6365  
 6370  
 6375  
 6380  
 6385  
 6390  
 6395  
 6400  
 6405  
 6410  
 6415  
 6420  
 6425  
 6430  
 6435  
 6440  
 6445  
 6450  
 6455  
 6460  
 6465  
 6470  
 6475  
 6480  
 6485  
 6490  
 6495  
 6500  
 6505  
 6510  
 6515  
 6520  
 6525  
 6530  
 6535  
 6540  
 6545  
 6550  
 6555  
 6560  
 6565  
 6570  
 6575  
 6580  
 6585  
 6590  
 6595  
 6600  
 6605  
 6610  
 6615  
 6620  
 6625  
 6



453

(A) NAME/KEY: peptide  
(B) LOCATION: 1..219  
(D) OTHER INFORMATION: / Ceres Seq. ID 1533355  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:271:  
Met Asp Gly Val Leu Cys Asn Ser Glu Asp Leu Ser Arg Arg Ala Ala  
1 5 10 15  
Val Asp Val Phe Thr Glu Met Gly Val Glu Val Thr Val Asp Asp Phe  
20 25 30  
Val Pro Phe Met Gly Thr Gly Glu Ala Lys Phe Leu Gly Gly Val Ala  
35 40 45  
Ser Val Lys Glu Val Lys Gly Phe Asp Pro Asp Ala Ala Lys Lys Arg  
50 55 60  
Phe Phe Glu Ile Tyr Leu Asp Lys Tyr Ala Lys Pro Glu Ser Gly Ile  
65 70 75 80  
Gly Phe Pro Gly Ala Leu Glu Leu Val Thr Glu Cys Lys Asn Lys Gly  
90 95  
Leu Lys Val Ala Val Ala Ser Ser Ala Asp Arg Ile Lys Val Asp Ala  
100 105 110  
Asn Leu Lys Ala Val Gly Leu Ser Leu Thr Met Phe Asp Ala Ile Val  
115 120 125  
Ser Ala Asp Ala Phe Glu Asn Leu Lys Pro Ala Pro Asp Ile Phe Leu  
130 135 140  
Ala Ala Ala Lys Ile Leu Gly Val Pro Thr Ser Glu Cys Val Val Ile  
145 150 155  
Glu Asp Ala Leu Ala Gly Val Glu Ala Ala Glu Ala Ala Asn Met Arg  
160 165 170  
Cys Ile Ala Val Lys Thr Thr Leu Ser Glu Ala Ile Leu Lys Asp Ala  
175 180 185  
Gly Pro Ser Met Ile Arg Asp Asp Ile Gly Asn Ile Ser Ile Asn Asp  
190 195 200  
Ile Leu Thr Gly Ser Asp Ser Thr Ser Met  
205 210 215

(2) INFORMATION FOR SEQ ID NO:272:  
(A) LENGTH: 1420 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(xi) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..1420  
(D) OTHER INFORMATION: / Ceres Seq. ID 1534544  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:272:  
aaagattctt ggtctcgtt ccagccact gaagattctt ggagttctt ttgaccacac  
120 acttcaatg ccgagctgtg tgaacgtca aactccatcg ctctccgagc aatacatatt  
ggagaagaa ggaagcaag acacaagtcg aaagctgtt gaagtgaaag agtggcaacc  
240 agagttact acacaagctg agaggttaa gacggagcaa gctaaaggag aatctctgt  
tgagagagc attctctag ttgaagaaa gctgaattt gctctgaat caacggaggt  
360 ggctctcag gctcc-gctg cagcggaaga caatgctga gacactctg ctgctgtga  
420 agaaataar gagaanaacg ctatggaaga agttgtgaa gaacacctg atgagatcaa  
480 gcttgagaa gctcc-gctt gactcccaa ggggttctt tttttttt ttttttgaat  
540 ttttttcaa agtgtcgtt ccagctatg tttgtctcc agtatgata tctccgtgac  
600 tgcagttgtt tcaaggagaa attgaagag agcgttagag ttgggaaaa ctcatgtgtt  
660 taggctaaa ggaagaactc aagcaatcat tgcctgtta catgggttg gggacaatgg  
720 ctcaagcttg tccagcttt ttgaacctt tcccttcaa aatacaaat ggttt-gccc  
780 ggtgtctct tctcaacaa taagttatt tgg-ggttt cctcccaag cttgtttga  
840 tcttgtagc atcaatgag atgagctga tgg-atggaa ggtatggag tgggtgtgc  
900 acatgtgca aattgttgt cgaatgagc tgc-gacatt aaataggtt ttgagagatt  
960 cagatgggt gggcgacat cctcatatt tgaacttt ttgtctctg gtaaaatagg  
1020 aaatggcat caacaccta tcaattaa gcaatcata ggttcaagc ggtgtctcc  
1080 ttgtgnaag acatgtgctg gcaactaga agaggaacag atcaagaacc gagtgcact

454

gttaccattt gttctcgtc atggaagaac tgatgatgtg ataccattca agtttgggga  
1140 gaaatttca cagctttgc ttcaaaagg gtttaagaag gtacattca aactttacag  
1200 tgcatttgt cacaacaaa tcccacaga gtggagag gtgtgacat ggttgacatc  
1260 cagcttgcg cctgaaggtt gactctctt atgagtagc ttgtgatga aaaccttca  
1320 actctgaga gtttgatga atggatgtt tcagatttc acaatgttt cattgnaa  
1380 ttgtgtaag acacattct ctatgaaca atctttcctt

(2) INFORMATION FOR SEQ ID NO:273:  
(A) LENGTH: 146 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(xi) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..146  
(D) OTHER INFORMATION: / Ceres Seq. ID 1534545  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:273:  
Lys Ile Leu Gly Leu Arg Ser Ser His Leu Lys Ile Leu Gly Ala Ser  
1 5 10 15  
Phe Asp His Thr Leu Ser Met Ala Asp Ala Val Asn Ala Glu Thr Pro  
20 25 30  
Ser Leu Ser Glu Glu Tyr His Leu Glu Lys Glu Val Lys Glu Asp Thr  
35 40 45  
Ser Ala Lys Pro Val Glu Val Lys Glu Val Ala Pro Glu Val Thr Thr  
50 55 60  
Gln Ala Glu Glu Val Lys Thr Glu Glu Ala Lys Glu Ser Pro Val  
65 70 75  
Glu Glu Ala Val Ser Val Val Glu Glu Lys Ser Glu Ser Ala Pro Glu  
80 85 90  
Ser Thr Glu Val Ala Ser Glu Ala Pro Ala Ala Ala Glu Asp Asn Ala  
100 105 110  
Glu Glu Thr Pro Ala Ala Glu Glu Asn Asn Asp Glu Asn Ala Ser  
115 120 125  
Glu Glu Val Ala Glu Glu Thr Pro Asp Glu Ile Lys Leu Glu Thr Ala  
130 135 140  
Pro Ala  
145

(2) INFORMATION FOR SEQ ID NO:274:  
(A) LENGTH: 252 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(xi) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..252  
(D) OTHER INFORMATION: / Ceres Seq. ID 1534546  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:274:  
Met Ser Ile Ser Gly Ala Ala Val Gly Ser Gly Arg Asn Leu Arg Arg  
1 5 10 15  
Ala Val Glu Phe Gly Lys Thr His Val Val Arg Pro Lys Gly Lys His  
20 25 30  
Gln Ala Thr Ile Val Trp Leu His Gly Leu Gly Asp Asn Gly Ser Ser  
35 40 45  
Trp Ser Gln Leu Leu Glu Thr Leu Pro Leu Pro Asn Ile Lys Trp Ile  
50 55 60  
Cys Pro Thr Ala Pro Ser Gln Pro Ile Ser Leu Phe Gly Gly Phe Pro  
65 70 75 80  
Ser Thr Ala Trp Phe Asp Val Val Asp Ile Asn Glu Asp Gly Pro Asp  
85 90

455

Asp Met Glu Gly Leu Asp Val Ala Ala His Val Ala Asn Leu Leu  
 100 105 110  
 Ser Asn Glu Pro Ala Asp Ile Lys Leu Gly Val Gly Gly Phe Ser Met  
 115 120 125  
 5 Gly Ala Ala Thr Ser Leu Tyr Ser Ala Thr Cys Phe Ala Leu Gly Lys  
 130 135 140  
 Tyr Gly Asn Gly Asn Pro Tyr Pro Ile Asn Leu Ser Ala Ile Ile Gly  
 145 150 155 160  
 Leu Ser Gly Trp Leu Pro Cys Ala Lys Thr Leu Ala Gly Lys Leu Glu  
 170 175  
 10 Glu Glu Gic Ile Lys Asn Arg Ala Ala Ser Leu Pro Ile Val Val Cys  
 180 185 190  
 His Gly Lys Ala Asp Val Val Pro Phe Lys Phe Gly G.C Lys Ser  
 195 200 205  
 15 Ser Gln Ala Leu Leu Ser Asn Gly Phe Lys Lys Val Thr Phe Lys Pro  
 210 215 220  
 Tyr Ser Ala Leu Gly His His Thr Ile Pro Gln Glu Leu Asp Glu Leu  
 225 230 235 240  
 Cys Ala Trp Leu Thr Ser Thr Leu Ser Leu Glu Gly  
 245 250

(2) INFORMATION FOR SEQ ID NO:275:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 155 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

(A) NAME/KEY: peptide  
 (B) LOCATION: 1..155

(D) OTHER INFORMATION: / Ceres Seq. ID 1534547

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:275:

Met Glu Gly Leu Asp Val Ala Ala His Val Ala Asn Leu Leu Ser  
 1 5 10 15  
 35 Asn Glu Pro Ala Asp Ile Lys Leu Gly Val Gly Gly Phe Ser Met Gly  
 20 25 30  
 Ala Ala Thr Ser Leu Tyr Ser Ala Thr Cys Phe Ala Leu Gly Lys Tyr  
 35 40 45  
 Gly Asn Gly Asn Pro Tyr Pro Ile Asn Leu Ser Ala Ile Ile Gly Leu  
 50 55 60  
 40 Ser Gly Trp Leu Pro Cys Ala Lys Thr Leu Ala Gly Lys Leu Glu Glu  
 65 70 75 80  
 Glu G.C Ile Lys Asn Arg Ala Ala Ser Leu Pro Ile Val Val Cys His  
 85 90 95  
 45 Gly Lys Ala Asp Asp Val Val Pro Phe Lys Phe Gly G.C Lys Ser Ser  
 100 105 110  
 Gln Ala Leu Leu Ser Asn Gly Phe Lys Lys Val Thr Phe Lys Pro Tyr  
 115 120 125  
 Ser Ala Leu Gly His Thr Ile Pro Gln Glu Leu Asp Glu Leu Cys  
 130 135 140  
 50 Ala Trp Leu Thr Ser Thr Leu Ser Leu Glu Gly  
 145 150 155

(2) INFORMATION FOR SEQ ID NO:276:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1592 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:

(A) NAME/KEY: -  
 (B) LOCATION: 1..1592

456

(D) OTHER INFORMATION: / Ceres Seq. ID 1567172  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:276:

gtatgcccc ttttataac cactctctga aaactgaac ctttgagag agaacccata  
 60  
 gtgataaa acattcttt tcaactgag acttgcaac ttgttttac tcaagttaag  
 120  
 attctcttt gtgatgggtg ataagttcaa tatgatgaa agttggcta gtgggttgg  
 180  
 aggaagaag cggagggaat gtgatagctg gtgaagagg cggagacgtt gattatcgc  
 240  
 actcgaatg gctttcttt gcatgtgttg tgaagttcgt gtcacatcgg caaacccct  
 300  
 tctcgttaa cacaagaag tctgttgaa atcgtgtgac cgcgaatac atcgcaagc  
 360  
 ctccgcctt caccaaqca ctggatcaa gggatttca cgtaaagtc ggaaccacg  
 420  
 10 ttagggaag aagagocaca ccatgtttt tcatgatct gtccggaga tgaacagg  
 480  
 ggaacagg gagaattac agtgaaga ggcctcata ttggaggtc cggatgaa  
 540  
 cctcatggt gaggagcaat gattaaaca atcccgag aacagaaatg agtttcaat  
 600  
 gatgcctta agtttcaaga gtatgacga agaatgac gacaacgtg aagattgtct  
 660  
 gaatggttg ttccaaacg acatggaact agtcaatc acagtgacg tggagactct  
 720  
 15 actcgggga ggggaacgag agtttcttc catagaaga ctagggttag tggagactct  
 780  
 aaagatgaa aaagaggag lggaggaga ggaaggagt gtgacaagag aagtgcata  
 840  
 tcaagatga ggtgatgaga catcccaat tgaataagc ttgatcac agtacaaca  
 900  
 caagacaca ttcatgag gagagaaga tgaagaaga gacgtgatg aagaatgac  
 960  
 20 ggaatgga gtgaatgga tgaatgttg gtttaagaa gagaagaag agatccctt  
 1020  
 tatgttga ttgactatg aatagcat ttcaattg ggaagcaag ggaatccatg  
 1080  
 ggcgcctg gtgcacttg aatagacct agatggtt tgttccaa ccataccat  
 1140  
 ggtgaaat ggaagagg ctcataca caaccatc cgcgcctag gttaccct  
 1200  
 aggaatgt gggatggag gaagaggag taggtttca agatccag aaaaaggag  
 1260  
 25 gacaagtg ttccaaga agataagta cgaagtcag aaatgaaat cagatcaag  
 1320  
 gctcgcag aaagaaagt tctcaagag acttcaatt ggtgtgct actaaagac  
 1380  
 taaatatt atgatatta aattacrtg ctcaattt gcttttgt tgcatagt  
 1440  
 ttgtgatg ttgattctt ttctctgc tcatagaga ttgtcagc ttgtgagc  
 1500  
 30 tactatga cataatara taccnaaa atgtagct ctgtgaaga ctgattata  
 1560

(2) INFORMATION FOR SEQ ID NO:277:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 407 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..407

(D) OTHER INFORMATION: / Ceres Seq. ID 1567173

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:277:

Met Met Lys Ser Leu Ala Ser Ala Val Gly Gly Lys Thr Ala Arg Ala  
 1 5 10 15  
 Cys Asp Ser Cys Val Lys Arg Arg Ala Arg Trp Tyr Cys Ala Ala Asp  
 20 25 30  
 45 Asp Ala Phe Leu Cys His Ala Cys Asp Gly Ser Val His Ser Ala Asn  
 35 40 45  
 Pro Leu Ala Arg Arg His Glu Arg Val Arg Leu Lys Ser Ala Ser Ala  
 50 55 60  
 50 Gly Lys Tyr Arg His Ala Ser Pro Pro His Gln Ala Thr Trp His Gln  
 65 70 75 80  
 Gly Phe Thr Arg Lys Ala Arg Thr Pro Arg Gly Gly Lys Lys Ser His  
 85 90 95  
 Thr Met Val Phe His Asp Leu Val Pro Glu Met Ser Thr Glu Asp Gln  
 100 105 110  
 55 Ala Glu Ser Tyr Glu Val Glu Glu Gln Leu Ile Phe Glu Val Pro Val  
 115 120 125  
 Met Asn Ser Met Val Glu Glu Gln Cys Phe Asn Gln Ser Leu Glu Lys  
 130 135 140  
 60 Gln Asn Glu Phe Pro Met Met Pro Leu Ser Phe Lys Ser Ser Asp Glu  
 145 150 155 160  
 Glu Asp Asp Asp Asn Ala Glu Ser Cys Leu Asn Gly Leu Phe Pro Thr

457  
170  
165  
175  
Asp Met Glu Leu Ala Gln Phe Thr Ala Asp Val Glu Thr Leu Leu Gly  
185 190  
Gly Gly Asp Arg Glu Phe His Ser Ile Glu Glu Leu Gly Leu Gly  
195 200 205  
Met Leu Lys Ile Glu Lys Glu Glu Val Glu Glu Glu Gly Val Val  
210 215 220  
Thr Arg Glu Val His Asp Gln Asp Glu Gly Asp Glu Thr Ser Pro Phe  
225 230 235 240  
Glu Ile Ser Phe Asp Tyr Glu Tyr Thr His Lys Thr Thr Phe Asp Glu  
245 250 255  
Gly Glu Glu Asp Glu Lys Glu Asp Val Met Lys Asn Val Met Glu Met  
260 265 270  
Gly Val Asn Glu Met Ser Gly Gly Ile Lys Glu Glu Lys Lys Lys  
275 280 285  
Ala Leu Met Leu Arg Leu Asp Tyr Glu Ser Val Ile Ser Thr Trp Gly  
290 295 300  
Gly Gln Gly Ile Pro Trp Thr Ala Arg Val Pro Ser Glu Ile Asp Leu  
305 310 315 320  
Asp Met Val Cys Phe Pro Thr His Thr Met Gly Glu Ser Gly Ala Glu  
325 330 335  
Ala His His Asn His Phe Arg Gly Leu Glu His Leu Gly Asp  
340 345 350  
Ala Gly Asp Gly Gly Arg Glu Ala Arg Val Ser Arg Tyr Arg Glu Lys  
355 360 365  
Arg Arg Thr Arg Leu Phe Ser Lys Lys Ile Arg Tyr Glu Val Arg Lys  
370 375 380  
Leu Asn Ala Asp Lys Arg Pro Arg Met Lys Gly Arg Phe Val Lys Arg  
385 390 395 400  
Ser Ser Ile Gly Val Ala His  
405

(2) INFORMATION FOR SEQ ID NO:278:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 406 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..406  
(D) OTHER INFORMATION: / Ceres Seq. ID 1567174  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:278:  
Met Lys Ser Leu Ala Ser Ala Val Gly Gly Lys Thr Ala Arg Ala Cys  
1 5 10 15  
Asp Ser Cys Val Lys Arg Arg Ala Arg Trp Tyr Cys Ala Ala Asp Asp  
20 25 30  
Ala Phe Leu Cys His Ala Cys Asp Gly Ser Val His Ser Ala Asn Pro  
35 40 45  
Leu Ala Arg Arg His Glu Arg Val Arg Leu Lys Ser Ala Ser Ala Gly  
50 55 60  
Lys Tyr Arg His Ala Ser Pro Pro His Gln Ala Thr Trp His Gln Gly  
65 70 75 80  
Phe Thr Arg Lys Ala Arg Thr Pro Arg Gly Gly Lys Lys Ser His Thr  
85 90 95  
Met Val Phe His Asp Leu Val Pro Glu Met Ser Thr Glu Asp Gln Ala  
100 105 110  
Glu Ser Tyr Glu Val Glu Glu Gln Leu Ile Phe Glu Val Pro Val Met  
115 120 125  
Asn Ser Met Val Glu Glu Gln Cys Phe Asn Gln Ser Leu Glu Lys Gln  
130 135 140  
Asn Glu Phe Pro Met Met Pro Leu Ser Phe Lys Ser Asp Glu Glu

458  
155  
160  
Asp Asp Asp Asn Ala Glu Ser Cys Leu Asn Gly Leu Phe Pro Thr Asp  
165 170 175  
Met Glu Leu Ala Gln Phe Thr Ala Asp Val Glu Thr Leu Leu Gly Gly  
180 185 190  
Gly Asp Arg Glu Phe His Ser Ile Glu Glu Leu Gly Leu Gly Glu Met  
195 200 205  
Leu Lys Ile Glu Lys Glu Glu Val Glu Glu Glu Gly Val Val Thr  
210 215 220  
Arg Glu Val His Asp Glu Asp Glu Gly Asp Glu Thr Ser Pro Phe Glu  
225 230 235 240  
Ile Ser Phe Asp Tyr Glu Tyr Thr His Lys Thr Thr Phe Asp Glu Gly  
245 250 255  
Glu Glu Asp Glu Lys Glu Asp Val Met Lys Asn Val Met Glu Met Gly  
260 265 270  
Val Asn Glu Met Ser Gly Gly Ile Lys Glu Glu Lys Lys Glu Lys Ala  
275 280 285  
Leu Met Leu Arg Leu Asp Tyr Glu Ser Val Ile Ser Thr Trp Gly Gly  
290 295 300  
Gln Gly Ile Pro Trp Thr Ala Arg Val Pro Ser Glu Ile Asp Leu Asp  
305 310 315 320  
Met Val Cys Phe Pro Thr His Thr Met Gly Glu Ser Gly Ala Glu Ala  
325 330 335  
His His His Asn His Phe Arg Gly Leu Gly Leu His Leu Gly Asp Ala  
340 345 350  
Gly Asp Gly Gly Arg Glu Ala Arg Val Ser Arg Tyr Arg Glu Lys Arg  
355 360 365  
Arg Thr Arg Leu Phe Ser Lys Lys Ile Arg Tyr Glu Val Arg Lys Leu  
370 375 380  
Asn Ala Asp Lys Arg Pro Arg Met Lys Gly Arg Phe Val Lys Arg Ser  
385 390 395 400  
Ser Ile Gly Val Ala His  
405

(2) INFORMATION FOR SEQ ID NO:279:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 310 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..310  
(D) OTHER INFORMATION: / Ceres Seq. ID 1567175  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:279:  
Met Val Phe His Asp Leu Val Pro Glu Met Ser Thr Glu Asp Gln Ala  
1 5 10 15  
Glu Ser Tyr Glu Val Glu Glu Gln Leu Ile Phe Glu Val Pro Val Met  
20 25 30  
Asn Ser Met Val Glu Glu Gln Cys Phe Asn Gln Ser Leu Glu Lys Gln  
35 40 45  
Asn Glu Phe Pro Met Met Pro Leu Ser Phe Lys Ser Ser Asp Glu Glu  
50 55 60  
Asp Asp Asp Asn Ala Glu Ser Cys Leu Asn Gly Leu Phe Pro Thr Asp  
65 70 75 80  
Met Glu Leu Ala Gln Phe Thr Thr Ala Asp Val Glu Thr Leu Leu Gly Gly  
85 90 95  
Gly Asp Arg Glu Phe His Ser Ile Glu Glu Leu Gly Leu Gly Glu Met  
100 105 110  
Leu Lys Ile Glu Lys Glu Glu Val Glu Glu Glu Gly Val Val Thr  
115 120 125  
Arg Glu Val His Asp Gln Asp Glu Gly Asp Glu Thr Ser Pro Phe Glu



130 Ile Ser Phe Asp Tyr Glu Tyr Thr His Lys Thr Phe Asp Glu Gly 140  
 145 Glu Glu Asp Glu Lys Glu Asp Val Met Lys Asn Val Met Glu Met Gly 150  
 155 Val Asn Glu Met Ser Gly Gly Ile Lys Lys Glu Lys Lys Lys Lys Ala 160  
 165 Leu Met Leu Arg Leu Asp Tyr Glu Ser Val Ile Ser Thr Trp Gly Gly 170  
 180 Gln Gly Ile Pro Trp Thr Ala Arg Val Pro Ser Glu Ile Asp Leu Asp 185  
 195 Met Val Cys Phe Pro Thr His Thr Met Gly Glu Ser Gly Ala Glu Ala 200  
 210 His His Asn His Phe Arg Gly Leu Gly Leu His Leu Gly Asp Ala 205  
 225 Gly Asp Gly Gly Arg Glu Ala Arg Val Ser Arg Tyr Arg Glu Lys Arg 210  
 230 Arg Thr Arg Leu Phe Ser Lys Lys Ile Arg Tyr Glu Val Arg Lys Leu 215  
 240 Asn Ala Asp Lys Arg Pro Arg Met Lys Gly Arg Phe Val Lys Arg Ser 220  
 250 Ser Ile Gly Val Ala His 225  
 260

(2) INFORMATION FOR SEQ ID NO:280:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 520 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..520  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1567536

(1) SEQUENCE DESCRIPTION: SEQ ID NO:280:

120 aaaaattccg cagaattgat tatctcatt tcttgaagga tctctgtggg gttgatcga  
 130 attctcagc gaagtagaga g-gagagaaga gattgaagca gtaatggcag gaacattgg  
 140 attgtcaac gcatgaagc caaagatcca aacgattgat attcaggccg ccgcggagag  
 150 gggaaatccc gccgagccg g-gccattcg ggcgtccca ccaattgggt gatacaagaa  
 160 gacattcatt gaccacccc caactgaaga gaagtagat atcaacaaga ttcaagacta  
 170 agtgaagagt ttccacattt caaagtgtg taacgcctct ttgcttga gctgaatata  
 180 ttgtgttct ttctcgggc t-tctgggtt cattctcca atgttgagt gattgtctt  
 190 cccaatqtt atacaaca agaaagac ttactccaq ttactgaa aagaaccca  
 200 ttgtgcatt atcatgatt attattcc atgtcaagt

(2) INFORMATION FOR SEQ ID NO:281:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 57 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..57  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1567536

(1) SEQUENCE DESCRIPTION: SEQ ID NO:281:

1 Met Ala Gly Thr Ser Gly Leu Leu Asn Ala Val Lys Pro Lys Ile Gln  
 2 Thr Ile Asp Ile Gln Ala Ala Ala Gly Trp Gly Ile Ala Ala Ala Ala  
 3 Gly Ala Ile Trp Val Val Gln Pro Phe Gly Trp Ile Lys Lys Thr Phe  
 4

Ile Asp Pro Pro Pro Thr Glu Glu Lys 460

50 55

(2) INFORMATION FOR SEQ ID NO:282:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 765 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..765  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1569689

(1) SEQUENCE DESCRIPTION: SEQ ID NO:282:

15 aacaaaaat ttctcgtct cgtctccaa aatcgatca aatctctaa aattcatt  
 20 ttctgttt gttctttt ttcttaag aatgcttca attctgca cttgcttc  
 25 tctactga ggtcagaatg a-ctttctc attactctt cctcdaaac tccatttc  
 30 cagctcag gtbaagcta g-ggagaag ccagatcca tgaactatc tgcactgt  
 35 tagacgatt cagaatggtt gggacagc tgaatagg ttagtctta ttggttgg  
 40 ttctgctt tatttagct cttgggcat cattgaat caccagca attgcaat  
 45 ttscgttt ctgagcgga t-cgaactg ttgtatctt gttctcac tggttcacat  
 50 atcgatat ctgttcaaa ccgacagac agagcttcc gaaat-gtc aagaatcag  
 55 tagcggtat acilggucag tgaacttgt gttgtgata atacttcat ctttgaga  
 60 ttattggtt gcaagtgtt a-aattcat gacaggggg ttgtgttc tagtcaata  
 65 agtcagta ttgaaacct gtaaatcull attgttgt ttgtgtgt gagaatac  
 70 atctttct aattcaag attctctt atgattatc ttttt

(2) INFORMATION FOR SEQ ID NO:283:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 66 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..66  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1569690

(1) SEQUENCE DESCRIPTION: SEQ ID NO:283:

40 Asn Gln Lys Phe Leu Arg Leu Cys Leu Gln Asn Arg Ile Lys Ile Ser  
 5  
 10  
 15  
 20 Lys Val Ser Ile Phe Leu Phe Cys Ser Phe Phe Phe Leu Lys Asn Gly  
 25  
 30 Phe Asn Phe Cys Asn Phe Ala Phe Ala Ile Val Thr His Thr Glu Lys  
 35  
 40 Ile Gln Ser His Ile Asp Ser Lys Thr Pro Ile Phe Ser Asn Ser Arg  
 45  
 50 Tyr Gln  
 55

(2) INFORMATION FOR SEQ ID NO:284:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 107 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..107  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1569691

(1) SEQUENCE DESCRIPTION: SEQ ID NO:284:

Met Ala Ser Ile Ser Ala Thr Leu Pro Ser Pro Leu Leu Thr Gln

1 Arg Lys Ser Asn Leu Thr Ser Ile Gln Lys Leu Pro Phe Ser Leu Thr 15  
20 30  
5 Arg Gly Thr Asn Asp Leu Ser Pro Leu Ser Leu Thr Arg Asn Pro Ser 45  
35 45  
Ser Ile Ser Leu Met Val Lys Ala Ser Gly Gln Ser Ser Asp Ser Ser 60  
50 60  
Thr Asp Leu Asp Val Val Ser Thr Ile Gln Asn Val Trp Asp Lys Ser 75  
65 75  
10 Glu Asp Arg Leu Gly Leu Ile Gly Leu Gly Phe Xaa Trp Tyr Cys Ser 95  
85 95  
Ser Leu Gly Ile Ile Glu Ser His His Gly Asn 100 105  
(2) INFORMATION FOR SEQ ID NO:285:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 55 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..55  
(D) OTHER INFORMATION: / Ceres Seq. ID 1569692  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:285:  
Met Val Lys Ala Ser Gly Glu Ser Ser Asp Ser Thr Asp Leu Asp 15  
1 10 15  
Val Val Ser Thr Ile Gln Asn Val Trp Asp Lys Ser Glu Asp Arg Leu 20  
25 30  
30 Gly Leu Ile Gly Leu Gly Phe Xaa Trp Tyr Cys Ser Ser Leu Gly Ile 45  
35 45  
Ile Glu Ser His His Gly Asn 50 55  
(2) INFORMATION FOR SEQ ID NO:286:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 782 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..782  
(D) OTHER INFORMATION: / Ceres Seq. ID 1571042  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:286:  
atcaaccgaa acaaacata aanaaagt ggaagctta aacagagg gaggagcga 60  
aatggcag gtcggagac tacygagc agttccgt aaaggatc ggcagagat 120  
actacagc gtagagac ggaacatg tcttcctg tgcctcgc caccatcc 180  
aaatgtac gttcacga gsgaacatg actctcac gttctcag agttggaat 240  
acacatgga tggagagg gagggttca aggagaga agatagac gtagacca 300  
aaccttgc gtaagaga ctggggct acgtgaga gcagtcga gtagagacy 360  
tgcttaca attatccc aaatcagg atactcgt cggaaaatc accatcat 420  
ggagaagc caacgatg tccagagc caagcgct catgaatt gtaagagt 480  
tggttctg catggaac cagctaga aacttaac atcatlcca cagctcgt 540  
caccatcc atcatcra tcatcatc atcatcgc atcatcgc tcatcatc 600  
caccatcc atcatcra tcatcatc atcatcgc atcatcgc atcatcgc 660  
gggtcttg gatcttcat tctatggt aaacgttg gttctgat atgtctgat 720  
atattgtat gttatgctc atatgggg tcatatata tgaattta gattattta 780  
(2) INFORMATION FOR SEQ ID NO:287:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 151 amino acids

461  
10 15  
Arg Lys Ser Asn Leu Thr Ser Ile Gln Lys Leu Pro Phe Ser Leu Thr 15  
20 30  
5 Arg Gly Thr Asn Asp Leu Ser Pro Leu Ser Leu Thr Arg Asn Pro Ser 45  
35 45  
Ser Ile Ser Leu Met Val Lys Ala Ser Gly Gln Ser Ser Asp Ser Ser 60  
50 60  
Thr Asp Leu Asp Val Val Ser Thr Ile Gln Asn Val Trp Asp Lys Ser 75  
65 75  
10 Glu Asp Arg Leu Gly Leu Ile Gly Leu Gly Phe Xaa Trp Tyr Cys Ser 95  
85 95  
Ser Leu Gly Ile Ile Glu Ser His His Gly Asn 100 105  
(2) INFORMATION FOR SEQ ID NO:288:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 718 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..718  
(D) OTHER INFORMATION: / Ceres Seq. ID 1571079  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:288:  
atcaagag acacaaata atgagttt ttttttaa ttgaaacaa atggggttga 60  
gtgggtct catgtggg gtggagta agtcocgg tgaagatc tggatgccc 120  
tcggagcg catcaatc tttccaaag attccctaa cgaactaaa accatcaag 180  
ttctagcgg gtagcgac gttctcgt ccat-cpcc: cacttctat gagaaggt 240  
ctccatggt gaagatcg gcgagaga tgaagcag: gatttggag acaaaaga 300  
tcctacag catctagg gcgaaact tggagtacta caaacgttc aaagaaac 360  
tgacgtat tcttaaga gtcgagcg tctgaaatg gtcgtgag tttagaaga 420  
ccgccatg attgaagc caacgtca tcaagact: tgcgtcaag aactcaag 480  
agatagat atactcct aagaacata gtcctaca ctgaacatt taattatat 540  
Maagaggtt gacatctc tataagatt tctaataa gaagttgat aaagtgaac 600  
ctcttata atcccaag tctgattc ggtttatg cagctaga ggcataagc 660  
ttttcaaa agccaatta gtcgaact ttgaaaaa: cgaactttt gtaaaagc 660  
(2) INFORMATION FOR SEQ ID NO:289:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 155 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide

463

(B) LOCATION: 1..155  
(D) OTHER INFORMATION: / Ceres Seq. ID 1572080  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:289:  
Met Gly Leu Ser Gly Val Leu His Val Glu Val Glu Val Lys Ser Pro  
1 5 10 15  
Ala Glu Lys Phe Thr Val Ala Leu Gly Asp Gly Ile Asn Leu Phe Pro  
20 25 30  
Lys Ala Phe Pro Asn Asp Tyr Lys Thr Ile Gln Val Leu Ala Gly Asp  
35 40 45  
Gly Asn Ala Pro Gly Ser Ile Arg Leu Ile Thr Tyr Gly Glu Gly Ser  
50 55 60  
Pro Leu Val Lys Ile Ser Ala Glu Arg Ile Glu Ala Val Asp Leu Glu  
65 70 75 80  
Asn Lys Ser Met Ser Tyr Ser Ile Ile Gly Gly Glu Met Leu Glu Tyr  
100 105 110  
Tyr Lys Thr Phe Lys Gly Thr Ile Thr Val Ile Pro Lys Asn Gly Gly  
115 120 125  
Ser Leu Leu Lys Tyr Ser Gly Glu Phe Glu Lys Thr Ala His Glu Ile  
130 135 140  
Asp Asp Pro His Val Ile Lys Asp Phe Ala Val Lys Asn Phe Lys Glu  
145 150 155  
Ile Asp Glu Tyr Leu Leu Lys Lys Glu Thr Ser Ala  
160 165 170 175 180 185 190 195 200

(2) INFORMATION FOR SEQ ID NO:290:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 808 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..808  
(D) OTHER INFORMATION: / Ceres Seq. ID 1572097  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:290:  
atcaaccaca acaaacacata aaagaacagt ggaagcttta aaacgagagg gataagacaa  
120  
aaatggcagc gtcgggaacg tacggagcgg aaattccgct aagaagatcg gcagagaatt  
180  
actacaaggc ggggaagac gagaacccatg tcttccctga tctatccg caccacatcc  
240  
aaatgttac cgttcacga ggcgaacatg actctccagg gtctatccag agttggact  
300  
acacatggga tgaagaagg gagggtttca agagagagag agagatagac gatgagacca  
360  
aaacgttac cttaagaga ctgaggttc acgtga-gaa gcagctcaaa gttacgaag  
420  
tgcttaccat attatccc aaattgagg atacctgcat cggcaaaatc acttaatat  
480  
gggagaagg caacgatgat tcccagac caagcggtca catgaattc gtcaagagt  
540  
tggttgctga catgggaac cagttgagcc ccaaaaaaaa aaaaaaaa cttatcatc  
600  
attccacag tctgctgat catcatcatc attcatcatc tcatcatcat catcatcatc  
660  
atcatcatc tcatcatcat catcatcatc atcatcatct cgattataa gtaagatgt  
720  
ctgtatgatg ctctgatata ttgtatgt ctatgatc ctgcatcctc tatgtgaaa cgtttggtt  
780  
tcttaagat taattacta caaattc

(2) INFORMATION FOR SEQ ID NO:291:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 195 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..195  
(D) OTHER INFORMATION: / Ceres Seq. ID 1572098  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:291:  
Met Ala Thr Ser Gly Thr Tyr Val Thr Glu Val Pro Leu Lys Gly Ser

464

Ala Glu Lys Tyr Tyr Lys Arg Tyr Lys Asn Glu Asn His Val Phe Pro  
1 5 10 15  
Asp Ala Ile Gly His His Ile Gln Asn Val Thr Val His Glu Gly Glu  
20 25 30  
His Asp Ser His Gly Ser Ile Arg Ser Trp Asn Tyr Thr Trp Asp Gly  
35 40 45  
Lys Glu Glu Val Phe Lys Glu Arg Arg Glu Ile Asp Asp Glu Thr Lys  
50 55 60  
Thr Leu Thr Leu Arg Gly Leu Glu Gly His Val Met Glu Gln Leu Lys  
65 70 75 80  
Val Tyr Asp Val Val Tyr Gln Phe Ile Pro Lys Ser Glu Asp Thr Cys  
85 90 95  
Ile Gly Lys Ile Thr Leu Ile Trp Glu Lys Arg Asn Asp Asp Ser Pro  
100 105 110  
Glu Pro Ser Gly Tyr Met Lys Phe Val Lys Ser Leu Val Ala Asp Met  
115 120 125  
Gly Asn His Val Ser Pro Lys Lys Lys Lys Lys Leu Asn His His  
130 135 140  
Ser His Ser Arg Arg His His His His His His His His His  
145 150 155 160  
His His His His His His His His His His His His His His  
165 170 175 180 185 190 195  
Leu Asp Leu

(2) INFORMATION FOR SEQ ID NO:292:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 104 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..104  
(D) OTHER INFORMATION: / Ceres Seq. ID 1572099  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:292:  
Met Glu Gln Leu Lys Val Tyr Asp Val Val Tyr Gln Phe Ile Pro Lys  
1 5 10 15  
Ser Glu Asp Thr Cys Ile Gly Lys Ile Thr Leu Ile Trp Glu Lys Arg  
20 25 30  
Asn Asp Asp Ser Pro Glu Pro Ser Gly Tyr Met Lys Phe Val Lys Ser  
35 40 45  
Leu Val Ala Asp Met Gly Asn His Val Ser Pro Lys Lys Lys Lys  
50 55 60  
Lys Leu Asn His His Ser His Ser Arg Arg His His His His His  
65 70 75 80  
His His His His His His His His His His His His His His  
85 90 95  
His His His His Tyr Leu Asp Leu

(2) INFORMATION FOR SEQ ID NO:293:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 583 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..583  
(D) OTHER INFORMATION: / Ceres Seq. ID 1572090



467

attgtgttc tttaagagtg ttgactctgg tgatcgagata actatctctt tgtgaagct  
tctatatat tgaagagca ttctctg

(2) INFORMATION FOR SEQ ID NO:298:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 171 amino acids  
(B) TYPE: amino acid

(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..171

(D) OTHER INFORMATION: / Ceres Seq. ID 1573607

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:298:

Met Ala Lys Leu Val Met Leu Val Leu Val Leu Cys Ile Leu Pro Ala Ile  
1 10 15

Ala Met Ala Ala Arg Arg Ser Asn Ile Gly Lys Asn Thr Met Val Val  
20 25 30

Gln Gly Ser Thr Tyr Cys Asp Ile Cys Lys Phe Gly Phe Glu Thr Pro  
35 40 45

Glu Ser Ser Tyr Phe Ile Pro Gly Ala Thr Val Lys Leu Ser Cys Lys  
50 55 60

Asp Arg Lys Thr Met Glu Glu Val Tyr Thr Asp Lys Ala Val Ser Asp  
65 70 75

Lys Glu Gly Lys Tyr Lys Phe Ile Val His Asp Asp His Arg Asp Gln  
80 85 90

Met Cys Asp Val Leu Leu Val Lys Ser Ser Asp Lys Thr Cys Ser Lys  
100 105 110

Ile Ser Val Gly Arg Glu Lys Ser Arg Val Ile Leu Asn His Tyr Ser  
115 120 125

Gly Ile Ala Ser Gln Ile Arg His Ala Asn Asn Met Gly Phe Glu Lys  
130 135 140

Glu Val Ser Asp Val Phe Cys Ser Ala Leu Phe Gln Lys Tyr Met Val  
145 150 155

Asp Glu Asp Glu Asp Asp Ile Lys Asn His Leu  
160 165 170

(2) INFORMATION FOR SEQ ID NO:299:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 166 amino acids  
(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..166

(D) OTHER INFORMATION: / Ceres Seq. ID 1573608

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:299:

Met Leu Leu Val Leu Cys Ile Leu Pro Ala Ile Ala Met Ala Ala Arg  
1 5 10 15

Arg Ser Asn Ile Gly Lys Asn Thr Met Val Val Gln Gly Ser Thr Tyr  
20 25 30

Cys Asp Ile Cys Lys Phe Gly Phe Glu Thr Pro Glu Ser Ser Tyr Phe  
35 40 45

Ile Pro Gly Ala Thr Val Lys Leu Ser Cys Lys Asp Arg Lys Thr Met  
50 55 60

Glu Glu Val Tyr Thr Asp Lys Ala Val Ser Asp Lys Glu Gly Lys Tyr  
65 70 75

Lys Phe Ile Val His Asp Asp His Arg Asp Gln Met Cys Asp Val Leu  
80 85 85

Leu Val Lys Ser Ser Asp Lys Thr Cys Ser Lys Ile Ser Val Gly Arg  
90 95 100

468

Glu Lys Ser Arg Val Ile Leu Asn His Tyr Ser Gly Ile Ala Ser Gln  
115 120 125

Ile Arg His Ala Asn Asn Met Gly Phe Glu Lys Glu Val Ser Asp Val  
130 135 140

Phe Cys Ser Ala Leu Phe Gln Lys Tyr Met Val Asp Glu Asp Glu Asp  
145 150 155

Asp Ile Lys Asn His Leu  
160 165

(2) INFORMATION FOR SEQ ID NO:300:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 154 amino acids  
(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..154

(D) OTHER INFORMATION: / Ceres Seq. ID 1573609

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:300:

Met Ala Ala Arg Arg Ser Asn Ile Gly Lys Asn Thr Met Val Val Gln  
1 5 10 15

Gly Ser Thr Tyr Cys Asp Ile Cys Lys Phe Gly Phe Glu Thr Pro Glu  
20 25 30

Ser Ser Tyr Phe Ile Pro Gly Ala Thr Val Lys Leu Ser Cys Lys Asp  
35 40 45

Arg Lys Thr Met Glu Glu Val Tyr Thr Asp Lys Ala Val Ser Asp Lys  
50 55 60

Glu Gly Lys Tyr Lys Phe Ile Val His Asp Asp His Arg Asp Gln Met  
65 70 75

Cys Asp Val Leu Leu Val Lys Ser Ser Asp Lys Thr Cys Ser Lys Ile  
80 85 85

Ser Val Gly Arg Glu Lys Ser Arg Val Ile Leu Asn His Tyr Ser Gly  
100 105 110

Ile Ala Ser Gln Ile Arg His Ala Asn Asn Met Gly Phe Glu Lys Glu  
115 120 125

Val Ser Asp Val Phe Cys Ser Ala Leu Phe Gln Lys Tyr Met Val Asp  
130 135 140

Glu Asp Glu Asp Asp Ile Lys Asn His Leu  
145 150

(2) INFORMATION FOR SEQ ID NO:301:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 704 base pairs  
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..704

(D) OTHER INFORMATION: / Ceres Seq. ID 1573861

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:301:

aaaaaacga atcaaaactt caaagttc aatttttg ttctgtctt tttttttt  
60

taagaatgg ctcaaatct tgaacttg ccttcgcat tttactcac acagagaaa  
120

tcaatctca catcgatca aaactccca tttctctaa ttctaggac gaagatctt  
180

ttctcattat ctcttactcg aaacctagc agcatcagtc tgaatgtaa agctagtgga  
240

gaagctcag attcaatcac tgaatcagc gtgttagta cgattcaga ttttgggat  
300

aggtctgaaa atagggtagg tttattggt ttgggtttg ctgattagt agctcttgg  
360

gcatactga atctcatcac ggcattcac aaatgcgag ttatctcag cggattcga  
420

ctagttgta tctgttttc cagtggttc acatacgt atctctgtt caaacggac  
480

agacagagc ttctgaaat tgcagagaa tcaagacgg atactctgg cAcagtgac  
540

cttggtggtg tgataact tcatcttgg aagatgattt gtttgcagt ttgtaaaatt  
600

acatgacag gtaggtgtg ttctatgccc aataatgcca tgcatttggaa acctgtgaat 660  
actttattgt tggttttgg ttgagacaa aatcaarttt ttct

(2) INFORMATION FOR SEQ ID NO:302:  
(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 179 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..179

(D) OTHER INFORMATION: / Ceres Seq. ID 1573862

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:302:

Met Ala Ser Ile Ser Ala Thr Leu Pro Ser Pro Leu Leu Thr Gln 15  
1 Arg Lys Ser Asn Leu Thr Ser Ile Gln Lys Leu Pro Phe Ser Leu Thr 30

Arg Gly Thr Asn Asp Leu Ser Pro Leu Ser Leu Thr Arg Asn Pro Ser 45  
20 Ser Ile Ser Leu Met Val Lys Ala Ser Gly Glu Ser Ser Asp Ser Ser 60

Thr Asp Leu Asp Val Val Ser Thr Ile Gln Asn Val Trp Asp Lys Ser 75  
65 Glu Asn Arg Leu Gly Leu Ile Gly Leu Gly Phe Ala Gly Ile Val Ala 90

Leu Trp Ala Ser Leu Asn Leu Ile Thr Ala Ile Asp Lys Leu Pro Val 105  
110

Ile Ser Ser Gly Phe Glu Leu Val Gly Ile Leu Phe Ser Thr Trp Phe 125  
30 Thr Tyr Arg Tyr Leu Leu Phe Lys Pro Asp Arg Gln Glu Leu Ser Lys 140

Ile Val Lys Lys Ser Val Ala Asp Ile Leu Gly Thr Val Asn Leu Val 155  
145 Cys Val Ile Ile Leu His Leu Trp Lys Met Ile Cys Leu Gln Val Cys 175

Lys Ile Thr 165

(2) INFORMATION FOR SEQ ID NO:303:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 127 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..127

(D) OTHER INFORMATION: / Ceres Seq. ID 1573863

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:303:

Met Val Lys Ala Ser Gly Glu Ser Ser Asp Ser Ser Thr Asp Leu Asp 15  
1 Val Val Ser Thr Ile Gln Asn Val Trp Asp Lys Ser Glu Asn Arg Leu 30

Gly Leu Ile Gly Leu Gly Phe Ala Gly Ile Val Ala Leu Trp Ala Ser 45  
55 Leu Asn Leu Ile Thr Ala Ile Asp Lys Leu Pro Val Ile Ser Ser Gly 60

Phe Glu Leu Val Gly Ile Leu Phe Ser Thr Trp Phe Thr Tyr Arg Tyr 75  
80

Leu Leu Phe Lys Pro Arg Gln Glu Leu Ser Lys Ile Val Lys Lys 95  
85

Ser Val Ala Asp Ile Leu Gly Thr Val Asn Leu Val Cys Val Ile Ile 110  
100

Leu His Leu Trp Lys Met Ile Cys Leu Gln Val Cys Lys Ile Thr 125  
115

5 (2) INFORMATION FOR SEQ ID NO:304:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 662 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..662

(D) OTHER INFORMATION: / Ceres Seq. ID 1574093

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:304:

attttttt tctctctc cctcttaagc aaaaataaaa caagctatgg ctggtatgct 60  
tcccgagttt aggtgtgcaa ggaaggggggg cttccacggtt ggtgctctc cgaatgaac 120

ctctctttac acttcaatc atgagtgcca ctaggcccac gttctcttct cgaagaaga 240  
-gttaggaat aaattttatg gagaagacaa cgaatgagaa cttgagggag cagcaaga 300

ggcaagaag aggttcaaa agtggctdag aatccacca gttacaagt cagcaaaat 360  
ggcaagaag aagggaaa aa attggagcaa ggaagggtta aactctcgg ggaattaccg 420

accgaagggtt ggcgggttaa agagagccg aggaagggtt atggaatggt tcaagcggcg 480  
agttaggaaa caacaagatt gtgtatatg tttagaccgg ttcaagaagg gtgagccctt 540

ggcacacca ccatgtgcc atagtttca cttcatatgc ttatgtcctt ggttagacac 600  
taagtattat tgcctatatt ggaagactga tattgggaa taagtattat atttttgatg 660

(2) INFORMATION FOR SEQ ID NO:305:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 198 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..198

(D) OTHER INFORMATION: / Ceres Seq. ID 1574094

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:305:

Met Ala Gly Met Leu Pro Gly Val Glu Cys Ala Arg Arg Arg Arg Phe 15  
1 His Gly Gly Ala Pro Ile Glu Ser Scr Asn Thr Ala Ser Val Ala 30

Ala Ala Ala Gly His Val Trp Thr Arg Arg Pro Ser Phe Ser Leu Tyr 45  
Thr Thr Asn His Glu Ser His Gln Ala His Val Ser Phe Ser Glu Arg 60

Ser Val Arg Asn Lys Ser Tyr Gly Glu Asp Asn Asp Glu Lys Leu Asp 80  
85 Ala Ala Lys Glu Ala Lys Gln Arg Leu Asn Lys Arg Leu Arg Ile 95

Pro Pro Arg Thr Ser Ser Gly Lys Met Val Lys Thr Lys Gly Ile Asn 110  
100

Trp Ser Lys Glu Arg Val Asn Leu Ser Gly Thr Tyr Arg Pro Arg Trp 125  
115

Val Gly Leu Lys Lys Ser Arg Gly Arg Leu Met Glu Trp Phe Lys Arg 140  
130

Arg Val Arg Glu Gln Asp Cys Ala Ile Cys Leu Asp Arg Phe Lys 160  
145

Lys Gly Glu Thr Leu Val His Leu Pro Cys Ala His Lys Phe His Ser 175  
165

471

Ile Cys Leu Leu Pro Trp Trp Leu Asp Thr Asn Val Tyr Cys Pro Tyr Cys  
180 185  
Arg Thr Asp Ile Trp Asn  
195

## 5 (2) INFORMATION FOR SEQ ID NO:306:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 195 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..195

(D) OTHER INFORMATION: / Ceres Seq. ID 1574095

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:306:

Met Leu Pro Gly Val Glu Cys Ala Arg Arg Arg Phe His Gly Gly  
1 10 15  
Ala Pro Pro Ile Glu Ser Ser Asn Thr Ala Ser Val Ala Ala Ala  
20 25 30  
Gly His Val Trp Thr Arg Arg Pro Ser Phe Ser Leu Tyr Thr Asn  
35 40 45  
His Glu Ser His Glu Ala His Val Ser Phe Ser Glu Arg Ser Val Arg  
50 55 60  
Asn Lys Ser Tyr Gly Glu Asp Asn Asp Glu Lys Leu Asp Gly Ala Ala  
65 70 75 80  
Lys Glu Ala Lys Glu Arg Leu Asn Lys Arg Leu Arg Ile Pro Pro Arg  
85 90 95  
Thr Ser Ser Gly Lys Met Val Lys Thr Lys Gly Ile Asn Trp Ser Lys  
100 105 110  
Glu Arg Val Asn Leu Ser Gly Thr Tyr Arg Pro Arg Trp Val Gly Leu  
115 120 125  
Lys Lys Ser Arg Gly Arg Leu Met Glu Trp Phe Lys Arg Arg Val Arg  
130 135 140  
Glu Glu Glu Asp Cys Ala Ile Cys Leu Asp Arg Phe Lys Lys Gly Glu  
145 150 155  
Thr Leu Val His Leu Pro Cys Ala His Lys Lys Phe His Ser Ile Cys Leu  
160 165 170  
Leu Pro Trp Leu Asp Thr Asn Val Tyr Cys Pro Tyr Cys Arg Thr Asp  
175 180 185  
Ile Trp Asn  
195

## (2) INFORMATION FOR SEQ ID NO:307:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 725 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..725

(D) OTHER INFORMATION: / Ceres Seq. ID 1580388

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:307:

ctcacacaa aatacaaaa cttagatcag tctcaagg ggaataaaa cttaaagaa  
60  
acataaag gcaacaata tcaacaata agcataaga agcataaga agcataaga  
120  
aagtagaa caagtgtac cagttaga gttctcttg tattgtatt gtaggttt  
160  
ctggtagc ctcgtgac tcatagcgg gcaacacat tcaagaag cgtatctc  
200  
ggaggaagt cagcgttcc aaacattcc accaacagg aaattcaaa acttgaag  
240  
tactcgtgg agcaattcaa tcaacaaga cagaacagg aggaacaa agtatccatt  
280  
gcgaacagg acagggcat ttcgaatca ttgcaatta gccgagtag gtcgtctag  
320  
aaacaggcgc tgcgtgact caataatat ctgaagattg aagtcactca acccaatggc  
360  
400  
480

472

tctacacaga tcttgactc tctgtggtt attcaacat ggtccattc taagcagttg  
540  
ctagttica cctcgtgtt cagctgttc tactaattt attctcttt attcagatta  
600  
aatccataa atagatcaa gaaagatta aaggtgtat gatacaaac tattagaat  
660  
gggttaalag ttggtttca tgatatgtt acgtgttca taataaaaa ctagttgta  
720  
ctagg

## (2) INFORMATION FOR SEQ ID NO:308:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 147 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..147

(D) OTHER INFORMATION: / Ceres Seq. ID 1580389

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:308:

Met Ala Thr Met Leu Lys Val Ser Leu Val Leu Ser Leu Leu Gly Phe  
1 5 10 15  
Leu Val Ile Ala Val Thr Pro Ser Ala Ala Asn Pro Phe Arg Lys  
20 25 30  
Ser Val Val Leu Gly Gly Lys Ser Gly Val Pro Asn Ile Arg Thr Asn  
35 40 45  
Arg Glu Ile Glu Glu Leu Gly Arg Tyr Cys Val Glu Glu Phe Asn Glu  
50 55 60  
Gln Ala Gln Asn Glu Glu Gly Asn Ile Gly Ser Ile Ala Lys Thr Asp  
65 70 75 80  
Thr Ala Ile Ser Asn Pro Leu Glu Phe Phe Sor Arg Val Val Ser Ala Gln  
85 90 95  
Lys Gln Val Val Ala Gly Leu Lys Tyr Leu Arg Ile Glu Val Thr  
100 105 110  
Gln Pro Asn Gly Ser Thr Arg Met Phe Asp Ser Val Val Ile Gln  
115 120 125  
Pro Trp Leu His Ser Lys Gln Leu Leu Gly Phe Thr Pro Val Val Ser  
130 135 140  
Pro Val Tyr  
145

## (2) INFORMATION FOR SEQ ID NO:309:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 144 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..144

(D) OTHER INFORMATION: / Ceres Seq. ID 1580390

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:309:

Met Leu Lys Val Ser Leu Val Leu Ser Leu Leu Gly Phe Leu Val Ile  
1 5 10 15  
Ala Val Val Thr Pro Ser Ala Ala Asn Pro Phe Arg Lys Ser Val Val  
20 25 30  
Leu Gly Gly Lys Ser Gly Val Pro Asn Ile Arg Thr Asn Arg Glu Ile  
35 40 45  
Gln Gln Leu Gly Arg Tyr Cys Val Glu Glu Phe Asn Gln Gln Ala Gln  
50 55 60  
Asn Glu Gln Gly Asn Ile Gly Ser Ile Ala Lys Thr Asp Thr Ala Ile  
65 70 75 80  
Ser Asn Pro Leu Gln Phe Ser Arg Val Val Ser Ala Gln Lys Gln Val  
85 90 95  
Val Ala Gly Leu Lys Tyr Tyr Leu Arg Ile Glu Val Thr Gln Pro Asn

473

100 105 110  
Gly Ser Thr Arg Met Phe Asp Ser Val Val Ile Gln Pro Trp Leu  
115 120 125  
His Ser Lys Gln Leu Leu Gly Phe Thr Pro Val Val Ser Pro Val Tyr  
130 135 140

5 (2) INFORMATION FOR SEQ ID NO:310:  
(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 675 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: DNA (genomic)

(1x) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..675

(3) OTHER INFORMATION: / Ceres Seq. ID 1582959

(1x1) SEQUENCE DESCRIPTION: SEQ ID NO:310:

tggcaggc acaaacata alcaactaa accttacta cttctattt ctttttaac  
60  
aaattacc tcttaacta tggcaggc gtttctact attctcttg tttttgctt  
120  
agctcagc ttacttgc ggcgaaga gttccagta gttctcttg tttttgctt  
180  
ctactcga tatgtggc gatttcgg cgttcggac aatggtttac ctttggcg  
240  
cgccggga gttgtctc gttccggg taacttgnl atygggat tgggtgtgc  
300  
tggcggc tggcggc gtttggcg tggcggcg tggcggcg atygggat  
360  
aggtggga agtgaatt gttcggac gttcggga agtccggg gtttctat  
420  
ccctggat gttacttg ttttaagg gtcacagg cttattaa gtttctat  
480  
cttaagat tttcataa ataatatc atactctt aggtttta acattggt  
540  
tatgaatt cttactgt ttaagtgc tttcaagta ctatttca cttgttga  
600  
atcattag tttgtgtc tgcattgc ttttctatt tcaattta cttgacct  
660  
ttcagttt tcaac

30 (2) INFORMATION FOR SEQ ID NO:311:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 115 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..115

(3) OTHER INFORMATION: / Ceres Seq. ID 1382960

(1x1) SEQUENCE DESCRIPTION: SEQ ID NO:311:

Met Ala Lys Trp Phe Thr Ile Phe Leu Val Phe Ala Leu Ala Ser  
1 10 15  
Ala Leu Ala Cys Gly Ala Arg Asn Val Pro Val Gly Leu Ser Asp Gln  
20 25 30  
Lys Asn Tyr Leu Gly Tyr Gly Gly Tyr Ser Gly Val Gly Asp Asn  
35 40 45  
Gly Leu Pro Phe Gly Gly Val Gly Gly Val Ser Gly Pro Gly Gly  
50 55 60  
Asn Leu Gly Tyr Gly Phe Gly Gly Ala Gly Gly Leu Gly Gly  
65 70 75  
Gly Leu Gly Gly Ala Gly Ser Gly Leu Gly Gly Gly Gly  
80 85 90 95  
Gly Ser Gly Ile Gly Ala Gly Thr Ser Gly Gly Thr Gly Gly Val  
100 105 110  
His Phe Pro  
115

(2) INFORMATION FOR SEQ ID NO:312:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1076 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

474

(D) TOPOLOGY: linear  
(11) MOLECULE TYPE: DNA (genomic)  
(1x) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..1076

(3) OTHER INFORMATION: / Ceres Seq. ID 1663221

(1x1) SEQUENCE DESCRIPTION: SEQ ID NO:312:

aaacagct taccgatt cttcgtcca acatttagg ttcagagat cgttgatt  
60  
tcacatcac agtatggc actcaatca gcaagaagc aagtttga gggagagtg  
120  
tatctacq tgaatgat gaggttcca caagagagc agcagagat ggtactctg  
180  
ggttgaggt taggttact caatgggga ctgagattat catcagatt actgtactc  
240  
agaattct tggtagaag qggagagaa ttaggaaat gacattctt gtcagaaga  
300  
gattcaatt tccagtgc agtttgagc tctatgctg aaggtttac acagaggtc  
360  
ctgtgcat tgcctgggt gattcttca gttacaagc tctcgtggt cttgtgttc  
420  
gcaggcctg ttatgtgt ttgagattg ttatggagc tggagttag ggtatgag  
480  
tcattgtg tggaaagtc cgtctgac gttcaagtc atgaagtc aagatgatt  
540  
acatggttc atctgtcaa ccaactaagc atacatcga tctgcagc agacatgtt  
600  
-gtcagaca ggggtgtcg ggaatcagg tgaatcatc ctttctcgg gacctcagg  
660  
gcaatcag accaagaca ccaatgctg atgttgatg cttctctg ccaagatg  
720  
atgtgtcta ctctgccc gctcagctg ctgctcagc tactctgtg caagaagtc  
780  
cactcaaac cgtagattac ctgagatga ttctccagc ggtctagaga agacttttt  
840  
tactactac caatgggat tgttctctt ttttatact tttctactt tgaccttc  
900  
tagctgtac tttatctc tttaacaga caagttttt cctatgtgt ttttaacct  
960  
atgcaagt tcttgaggt tgaatgctt tggttcagc atttttgt ttttaagct  
1020  
cgtacaaga aagaagatga cgaacacat gcttglatr tgaatgtgt acttgg

(2) INFORMATION FOR SEQ ID NO:313:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 250 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..250

(3) OTHER INFORMATION: / Ceres Seq. ID 1663222

(1x1) SEQUENCE DESCRIPTION: SEQ ID NO:313:

Met Ala Thr Gln Ile Ser Lys Lys Arg Lys Phe Val Ala Asp Gly Val  
1 5 10 15  
Phe Tyr Ala Glu Leu Asn Glu Val Leu Thr Arg Glu Leu Ala Glu Asp  
20 25 30  
Gly Tyr Ser Gly Val Glu Val Arg Val Thr Pro Met Arg Thr Glu Ile  
35 40 45  
Ile Ile Arg Ala Thr Arg Thr Gln Asn Val Leu Gly Glu Lys Gly Arg  
50 55 60  
Arg Ile Arg Glu Leu Thr Ser Leu Val Glc Lys Arg Phe Lys Phe Pro  
65 70 75  
Val Asp Ser Val Glu Leu Tyr Ala Glu Lys Val Asn Asn Arg Gly Leu  
80 85 90 95  
Cys Ala Ile Ala Gln Ala Glu Ser Leu Arg Tyr Lys Leu Leu Gly Gly  
100 105 110  
Leu Ala Val Arg Arg Ala Cys Tyr Gly Val Leu Arg Phe Val Met Glu  
115 120 125  
Ser Gly Ala Lys Gly Cys Glu Val Ile Val Ser Gly Lys Leu Arg Ala  
130 135 140  
Ala Arg Ala Lys Ser Met Lys Phe Lys Asp Gly Tyr Met Val Ser Ser  
145 150 155  
Gly Gln Pro Thr Lys Glu Tyr Ile Asp Ala Val Arg His Val Leu  
160 165 170 175  
Leu Arg Gln Gly Val Leu Gly Ile Lys Val Lys Ile Met Leu Asp Trp  
180 185 190  
Asp Pro Thr Gly Lys Ser Gly Pro Lys Thr Pro Leu Pro Asp Val Val



475

195 Ile Ile His Ala Pro Lys Asp Val Val Tyr Ser Ala Pro Ala Gln  
200  
210 Ala Ala Pro Val Thr Leu Val Gln Glu Ala Pro Leu Thr Thr Val  
225  
230 Asp Tyr Pro Glu Met Ile Pro Pro Val Ala  
245  
250

(2) INFORMATION FOR SEQ ID NO:314:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 207 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..207

(D) OTHER INFORMATION: / Ceres Seq. ID 1663223

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:314:

1 Met Arg Thr Glu Ile Ile Ile Arg Ala Thr Arg Thr Gln Asn Val Leu  
5  
10 Gly Glu Lys Gly Arg Arg Ile Arg Glu Leu Thr Ser Leu Val Gln Lys  
20  
25 Arg Phe Lys Phe Pro Val Asp Ser Val Glu Leu Tyr Ala Glu Lys Val  
35  
40 Asn Asn Arg Gly Leu Cys Ala Ile Ala Gln Ala Glu Ser Leu Arg Tyr  
50  
55 Lys Leu Leu Gly Gly Leu Ala Val Arg Arg Ala Cys Tyr Gly Val Leu  
65  
70 Arg Phe Val Met Glu Ser Gly Ala Lys Gly Cys Glu Val Ile Val Ser  
85  
90 Gly Lys Leu Arg Ala Ala Arg Ala Lys Ser Met Lys Phe Lys Asp Gly  
100  
105 Tyr Met Val Ser Ser Gly Gln Pro Thr Lys Glu Tyr Ile Asp Ala Ala  
120  
125 Val Arg His Val Leu Leu Arg Gln Gly Val Leu Gly Ile Lys Val Lys  
140  
145 Ile Met Leu Asp Thr Asp Pro Thr Gly Lys Ser Gly Pro Lys Thr Pro  
155  
160 Leu Pro Asp Val Val Ile Ile His Ala Pro Lys Asp Val Val Tyr  
175  
180 Ser Ala Pro Ala Gln Ala Ala Ala Pro Val Thr Leu Val Gln Glu Ala  
185  
190 Pro Leu Thr Thr Val Asp Tyr Pro Glu Met Ile Pro Pro Val Ala  
205

(2) INFORMATION FOR SEQ ID NO:315:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 488 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..488

(D) OTHER INFORMATION: / Ceres Seq. ID 1663275

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:315:

60 gtgactaat tctcagggt ttgaagatc tctctcgt cgtttgta tctctcgtg  
120 agcaataga ggaatggt cacttaagt tatgaact tcatcgaa agtactggtc  
180 ctggatctg ttatgcgt gtgicgga actgcacg gattgattc attagtacc  
240 tgaactggt cagacagt tctctaga acgttaaga gattgattc attagtacc  
300 gttactaag caccacactc atgattgat cttatgata taacataga ggcgtcgtg

476

5 ggtatgctt ttaagtttt gtatgtttt aaattttt ttgtgaac cattgtatt  
360  
420 ttggagta ataatgtt gaaacctat taagatgct tratttga taaatgagt  
480  
tattgcat ctegtattt tgtttatca gcaaatgtt actaagaaga atttgcattg  
tattgac

(2) INFORMATION FOR SEQ ID NO:316:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 56 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..56

(D) OTHER INFORMATION: / Ceres Seq. ID 1663276

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:316:

1 Met Gly His Ser Asn Val Thr Asn Ser His Pro Lys Lys Tyr Gly Pro  
5  
10 Gly Ser Arg Leu Cys Arg Val Cys Gly Asn Ser His Gly Leu Ile Arg  
20  
25 Lys Tyr Gly Leu Asn Cys Cys Arg Gln Cys Phe Arg Ser Asn Ala Lys  
35  
40 Glu Ile Gly Phe Ile Lys Tyr Arg  
45

(2) INFORMATION FOR SEQ ID NO:317:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 38 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..38

(D) OTHER INFORMATION: / Ceres Seq. ID 1663277

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:317:

1 Met Glu Leu Ser Ser Glu Glu Val Arg Ser Thr Ile Ser Phe Met Pro  
5  
10 Cys Val Arg Glu Leu Ala Arg Ser Asp Pro Glu Val Thr Phe Glu Leu  
20  
25 Leu Gln Thr Val Phe Pro  
30

(2) INFORMATION FOR SEQ ID NO:318:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 759 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..759

(D) OTHER INFORMATION: / Ceres Seq. ID 1665304

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:318:

55 atcaacaaa acaaacata aaaaacagt ggaagcttta aaacgaggg gagagagcaa  
120 aaagcgagc gtccggaacg tacgtgacg aggttcgct aaagatggt gccgacaaat  
180 actcaaggg gtggaagaa gagaaacatg tcttcctga tctatcgg caccacatc  
240 aaatgttac cgttcacaa ggcgaacatg acttcacgg gctatcagg agttggaact  
300 acaatggga tgaagaagag gaggtgttca agagagaag agatagac gatgagcca  
360 aaacgttgc gttcaagga cttgaggtc acgtatgga gcagctcaa gtgtacgag  
420 tgcgtacca attactccc aaatcgaag atactcgcat cggcaaaact actttaat  
480 gggagaagcg caacgatgat tcccagaac caagcgata catgaattc gtcaagagct

- 477  
 5 tgggttcgca catgggaac cagcttasca aaactaatc atcattecca cagctgctgt 540  
 cytgccatc atcatcata tcatcatat cctcatcatc atcatcata tcatcatcat 600  
 cctcatcatc tcatcatata agttaagatg ttctgcatat ataataagg gctctgggga 660  
 tctgtatct ctatgiglaa accgttgtgt tctgtatgat gctctcatat attgttatgt 720  
 tcatgatcat atgtcgggtt cgtataatg attcttaag  
 (2) INFORMATION FOR SEQ ID NO:319:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 151 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ix) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..15:  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1665305  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:319:  
 Met Ala Thr Ser Gly Thr Tyr Val Thr Glu Val Pro Leu Lys Gly Ser 15  
 10 Ala Glu Lys Tyr Tyr Lys Arg Trp Lys Asn Glu Asn His Val Phe Pro 30  
 Asp Ala Ile Gly His His Ile Gln Asn Val Thr Val His Glu Gly Glu 45  
 His Asp Ser His Gly Ser Ile Arg Ser Trp Asn Tyr Thr Trp Asp Gly 60  
 Lys Glu Glu Val Phe Lys Glu Arg Arg Glu Ile Asp Asp Glu Thr Lys 75  
 Thr Leu Thr Leu Arg Gly Leu Glu Gly His Val Met Glu Gln Leu Lys 90  
 Val Tyr Asp Val Val Tyr Gln Phe Ile Pro Lys Ser Glu Asp Thr Cys 105  
 Ile Gly Lys Ile Thr Leu Ile Trp Glu Lys Arg Asn Asp Asp Ser Pro 120  
 Glu Pro Ser Gly Tyr Met Lys Phe Val Lys Ser Leu Val Ala Asp Met 135  
 Gly Asn His Val Ser Lys Thr 150  
 (2) INFORMATION FOR SEQ ID NO:320:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 602 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ix) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..602  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1709970  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:320:  
 50 gaagctctt atataaat atcagagaga taagggttt ctagttagt gttccggga 60  
 gacgcttag cagaggaag agagagaga agagagggg cagctactac catggtctg 120  
 ggacggagct ttgtcttct agtgagcaga agattatccc aggaagagg attaggttta 180  
 tccgactga ttctcaggt ttctgttct ttaactcaca atgtaaggg tacttcata 240  
 acaagttaa gccacccaag ctgualgga ctgccatga cagaagcaa cncagaagg 300  
 atgcacaca agagctctg agagagaga gagcgcac caagaagca tactcaaggt 360  
 catgtgttg tgcactctg tgaacttc agakagag agctgagag cctgaagtc 420  
 gtgacgagc cagggaagt gctcgcgclg agatcaagga aagaatcaca agacccaag 480  
 atgaagaa ggcacaag gtggaattg ctctcaagca acagaagtc aggtctaat 540  
 tcccaaac tctctctga tccaaggtc ctaaggttgc aggtgtgtgt ggcacaagct 600  
 (2) INFORMATION FOR SEQ ID NO:321:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 108 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ix) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..108  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1709971  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:321:  
 Met Tyr Arg Lys Gln Xaa Lys Lys Asp Ala Ala Gln Glu Ala Val Lys 15  
 10 Arg Arg Arg Arg Ala Thr Lys Lys Pro Tyr Ser Arg Ser Ile Val Gly 30  
 15 Ala Thr Leu Glu Val Ile Gln Lys Lys Arg Ala Glu Lys Pro Glu Val 45  
 Arg Asp Ala Ala Arg Glu Ala Ala Leu Arg Glu Ile Lys Glu Arg Ile 60  
 Lys Lys Thr Lys Asp Glu Lys Lys Lys Lys Lys Val Glu Phe Ala Ser 75  
 20 Lys Gln Gln Lys Val Lys Ala Asn Phe Pro Lys Ala Ala Ala Ser 90  
 Lys Gly Pro Lys Val Gly Gly Gly Gly Lys Arg 105  
 (2) INFORMATION FOR SEQ ID NO:322:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 59 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ix) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..59  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1709972  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:322:  
 Met Gln Pro Gly Lys Leu Leu Cys Val Arg Ser Arg Lys Glu Ser Lys 15  
 10 Arg Pro Lys Met Lys Arg Arg Leu Arg Arg Trp Asn Leu Leu Ser 25  
 20 Asn Arg Arg Ser Arg Leu Ile Ser Pro Lys Leu Leu His Pro Arg 45  
 Val Leu Arg Trp Glu Val Val Val Ala Asn Ala 55  
 (2) INFORMATION FOR SEQ ID NO:323:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 726 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ix) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..726  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1711273  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:323:  
 atctttgt ctctctcca ttcttctta ttatccaga acgaagaaa accatgaaaa 60  
 cagctgaaga aagaataaca caagagagc catggcggg attgaccda ttactagg 120  
 ttggagaca gtgtgtatcc gaagagagc tctatcact cagcttaagc gcaagcaga 180  
 gactg-caac gccctgtct gaagcggc cgtatttag accgtcgaa aatcatgc 240  
 tggatgaac aggtgtcat caagggcac ctcttgac acagaagag tagataga 300  
 -accgagaa ctatctcatg atcgtgtcc cactgtatt aagaagcca tcatgcaag 360

479  
 1 420  
 2 480  
 3 540  
 4 600  
 5 660  
 6 720  
 7 780  
 8 840  
 9 900  
 10 960  
 11 1020  
 12 1080  
 13 1140  
 14 1200  
 15 1260  
 16 1320  
 17 1380  
 18 1440  
 19 1500  
 20 1560  
 21 1620  
 22 1680  
 23 1740  
 24 1800  
 25 1860  
 26 1920  
 27 1980  
 28 2040  
 29 2100  
 30 2160  
 31 2220  
 32 2280  
 33 2340  
 34 2400  
 35 2460  
 36 2520  
 37 2580  
 38 2640  
 39 2700  
 40 2760  
 41 2820  
 42 2880  
 43 2940  
 44 3000  
 45 3060  
 46 3120  
 47 3180  
 48 3240  
 49 3300  
 50 3360  
 51 3420  
 52 3480  
 53 3540  
 54 3600  
 55 3660  
 56 3720  
 57 3780  
 58 3840  
 59 3900  
 60 3960  
 61 4020  
 62 4080  
 63 4140  
 64 4200  
 65 4260  
 66 4320  
 67 4380  
 68 4440  
 69 4500  
 70 4560  
 71 4620  
 72 4680  
 73 4740  
 74 4800  
 75 4860  
 76 4920  
 77 4980  
 78 5040  
 79 5100  
 80 5160  
 81 5220  
 82 5280  
 83 5340  
 84 5400  
 85 5460  
 86 5520  
 87 5580  
 88 5640  
 89 5700  
 90 5760  
 91 5820  
 92 5880  
 93 5940  
 94 6000  
 95 6060  
 96 6120  
 97 6180  
 98 6240  
 99 6300  
 100 6360  
 101 6420  
 102 6480  
 103 6540  
 104 6600  
 105 6660  
 106 6720  
 107 6780  
 108 6840  
 109 6900  
 110 6960  
 111 7020  
 112 7080  
 113 7140  
 114 7200  
 115 7260  
 116 7320  
 117 7380  
 118 7440  
 119 7500  
 120 7560  
 121 7620  
 122 7680  
 123 7740  
 124 7800  
 125 7860  
 126 7920  
 127 7980  
 128 8040  
 129 8100  
 130 8160  
 131 8220  
 132 8280  
 133 8340  
 134 8400  
 135 8460  
 136 8520  
 137 8580  
 138 8640  
 139 8700  
 140 8760  
 141 8820  
 142 8880  
 143 8940  
 144 9000  
 145 9060  
 146 9120  
 147 9180  
 148 9240  
 149 9300  
 150 9360  
 151 9420  
 152 9480  
 153 9540  
 154 9600  
 155 9660  
 156 9720  
 157 9780  
 158 9840  
 159 9900  
 160 9960  
 161 10020  
 162 10080  
 163 10140  
 164 10200  
 165 10260  
 166 10320  
 167 10380  
 168 10440  
 169 10500  
 170 10560  
 171 10620  
 172 10680  
 173 10740  
 174 10800  
 175 10860  
 176 10920  
 177 10980  
 178 11040  
 179 11100  
 180 11160  
 181 11220  
 182 11280  
 183 11340  
 184 11400  
 185 11460  
 186 11520  
 187 11580  
 188 11640  
 189 11700  
 190 11760  
 191 11820  
 192 11880  
 193 11940  
 194 12000  
 195 12060  
 196 12120  
 197 12180  
 198 12240  
 199 12300  
 200 12360  
 201 12420  
 202 12480  
 203 12540  
 204 12600  
 205 12660  
 206 12720  
 207 12780  
 208 12840  
 209 12900  
 210 12960  
 211 13020  
 212 13080  
 213 13140  
 214 13200  
 215 13260  
 216 13320  
 217 13380  
 218 13440  
 219 13500  
 220 13560  
 221 13620  
 222 13680  
 223 13740  
 224 13800  
 225 13860  
 226 13920  
 227 13980  
 228 14040  
 229 14100  
 230 14160  
 231 14220  
 232 14280  
 233 14340  
 234 14400  
 235 14460  
 236 14520  
 237 14580  
 238 14640  
 239 14700  
 240 14760  
 241 14820  
 242 14880  
 243 14940  
 244 15000  
 245 15060  
 246 15120  
 247 15180  
 248 15240  
 249 15300  
 250 15360  
 251 15420  
 252 15480  
 253 15540  
 254 15600  
 255 15660  
 256 15720  
 257 15780  
 258 15840  
 259 15900  
 260 15960  
 261 16020  
 262 16080  
 263 16140  
 264 16200  
 265 16260  
 266 16320  
 267 16380  
 268 16440  
 269 16500  
 270 16560  
 271 16620  
 272 16680  
 273 16740  
 274 16800  
 275 16860  
 276 16920  
 277 16980  
 278 17040  
 279 17100  
 280 17160  
 281 17220  
 282 17280  
 283 17340  
 284 17400  
 285 17460  
 286 17520  
 287 17580  
 288 17640  
 289 17700  
 290 17760  
 291 17820  
 292 17880  
 293 17940  
 294 18000  
 295 18060  
 296 18120  
 297 18180  
 298 18240  
 299 18300  
 300 18360  
 301 18420  
 302 18480  
 303 18540  
 304 18600  
 305 18660  
 306 18720  
 307 18780  
 308 18840  
 309 18900  
 310 18960  
 311 19020  
 312 19080  
 313 19140  
 314 19200  
 315 19260  
 316 19320  
 317 19380  
 318 19440  
 319 19500  
 320 19560  
 321 19620  
 322 19680  
 323 19740  
 324 19800  
 325 19860  
 326 19920  
 327 19980  
 328 20040  
 329 20100  
 330 20160  
 331 20220  
 332 20280  
 333 20340  
 334 20400  
 335 20460  
 336 20520  
 337 20580  
 338 20640  
 339 20700  
 340 20760  
 341 20820  
 342 20880  
 343 20940  
 344 21000  
 345 21060  
 346 21120  
 347 21180  
 348 21240  
 349 21300  
 350 21360  
 351 21420  
 352 21480  
 353 21540  
 354 21600  
 355 21660  
 356 21720  
 357 21780  
 358 21840  
 359 21900  
 360 21960  
 361 22020  
 362 22080  
 363 22140  
 364 22200  
 365 22260  
 366 22320  
 367 22380  
 368 22440  
 369 22500  
 370 22560  
 371 22620  
 372 22680  
 373 22740  
 374 22800  
 375 22860  
 376 22920  
 377 22980  
 378 23040  
 379 23100  
 380 23160  
 381 23220  
 382 23280  
 383 23340  
 384 23400  
 385 23460  
 386 23520  
 387 23580  
 388 23640  
 389 23700  
 390 23760  
 391 23820  
 392 23880  
 393 23940  
 394 24000  
 395 24060  
 396 24120  
 397 24180  
 398 24240  
 399 24300  
 400 24360  
 401 24420  
 402 24480  
 403 24540  
 404 24600  
 405 24660  
 406 24720  
 407 24780  
 408 24840  
 409 24900  
 410 24960  
 411 25020  
 412 25080  
 413 25140  
 414 25200  
 415 25260  
 416 25320  
 417 25380  
 418 25440  
 419 25500  
 420 25560  
 421 25620  
 422 25680  
 423 25740  
 424 25800  
 425 25860  
 426 25920  
 427 25980  
 428 26040  
 429 26100  
 430 26160  
 431 26220  
 432 26280  
 433 26340  
 434 26400  
 435 26460  
 436 26520  
 437 26580  
 438 26640  
 439 26700  
 440 26760  
 441 26820  
 442 26880  
 443 26940  
 444 27000  
 445 27060  
 446 27120  
 447 27180  
 448 27240  
 449 27300  
 450 27360  
 451 27420  
 452 27480  
 453 27540  
 454 27600  
 455 27660  
 456 27720  
 457 27780  
 458 27840  
 459 27900  
 460 27960  
 461 28020  
 462 28080  
 463 28140  
 464 28200  
 465 28260  
 466 28320  
 467 28380  
 468 28440  
 469 28500  
 470 28560  
 471 28620  
 472 28680  
 473 28740  
 474 28800  
 475 28860  
 476 28920  
 477 28980  
 478 29040  
 479 29100  
 480 29160  
 481 29220  
 482 29280  
 483 29340  
 484 29400  
 485 29460  
 486 29520  
 487 29580  
 488 29640  
 489 29700  
 490 29760  
 491 29820  
 492 29880  
 493 29940  
 494 30000  
 495 30060  
 496 30120  
 497 30180  
 498 30240  
 499 30300  
 500 30360  
 501 30420  
 502 30480  
 503 30540  
 504 30600  
 505 30660  
 506 30720  
 507 30780  
 508 30840  
 509 30900  
 510 30960  
 511 31020  
 512 31080  
 513 31140  
 514 31200  
 515 31260  
 516 31320  
 517 31380  
 518 31440  
 519 31500  
 520 31560  
 521 31620  
 522 31680  
 523 31740  
 524 31800  
 525 31860  
 526 31920  
 527 31980  
 528 32040  
 529 32100  
 530 32160  
 531 32220  
 532 32280  
 533 32340  
 534 32400  
 535 32460  
 536 32520  
 537 32580  
 538 32640  
 539 32700  
 540 32760  
 541 32820  
 542 32880  
 543 32940  
 544 33000  
 545 33060  
 546 33120  
 547 33180  
 548 33240  
 549 33300  
 550 33360  
 551 33420  
 552 33480  
 553 33540  
 554 33600  
 555 33660  
 556 33720  
 557 33780  
 558 33840  
 559 33900  
 560 33960  
 561 34020  
 562 34080  
 563 34140  
 564 34200  
 565 34260  
 566 34320  
 567 34380  
 568 34440  
 569 34500  
 570 34560  
 571 34620  
 572 34680  
 573 34740  
 574 34800  
 575 34860  
 576 34920  
 577 34980  
 578 35040  
 579 35100  
 580 35160  
 581 35220  
 582 35280  
 583 35340  
 584 35400  
 585 35460  
 586 35520  
 587 35580  
 588 35640  
 589 35700  
 590 35760  
 591 35820  
 592 35880  
 593 35940  
 594 36000  
 595 36060  
 596 36120  
 597 36180  
 598 36240  
 599 36300  
 600 36360  
 601 36420  
 602 36480  
 603 36540  
 604 36600  
 605 36660  
 606 36720  
 607 36780  
 608 36840  
 609 36900  
 610 36960  
 611 37020  
 612 37080  
 613 37140  
 614 37200  
 615 37260  
 616 37320  
 617 37380  
 618 37440  
 619 37500  
 620 37560  
 621 37620  
 622 37680  
 623 37740  
 624 37800  
 625 37860  
 626 37920  
 627 37980  
 628 38040  
 629 38100  
 630 38160  
 631 38220  
 632 38280  
 633 38340  
 634 38400  
 635 38460  
 636 38520  
 637 38580  
 638 38640  
 639 38700  
 640 38760  
 641 38820  
 642 38880  
 643 38940  
 644 39000  
 645 39060  
 646 39120  
 647 39180  
 648 39240  
 649 39300  
 650 39360  
 651 39420  
 652 39480  
 653 39540  
 654 39600  
 655 39660  
 656 39720  
 657 39780  
 658 39840  
 659 39900  
 660 39960  
 661 40020  
 662 40080  
 663 40140  
 664 40200  
 665 40260  
 666 40320  
 667 40380  
 668 40440  
 669 40500  
 670 40560  
 671 40620  
 672 40680  
 673 40740  
 674 40800  
 675 40860  
 676 40920  
 677 40980  
 678 41040  
 679 41100  
 680 41160  
 681 41220  
 682 41280  
 683 41340  
 684 41400  
 685 41460  
 686 41520  
 687 41580  
 688 41640  
 689 41700  
 690 41760  
 691 41820  
 692 41880  
 693 41940  
 694 42000  
 695 42060  
 696 42120  
 697 42180  
 698 42240  
 699 42300  
 700 42360  
 701 42420  
 702 42480  
 703 42540  
 704 42600  
 705 42660  
 706 42720  
 707 42780  
 708 42840  
 709 42900  
 710 42960  
 711 43020  
 712 43080  
 713 43140  
 714 43200  
 715 43260  
 716 43320  
 717 43380  
 718 43440  
 719 43500  
 720 43560  
 721 43620  
 722 43680  
 723 43740  
 724 43800  
 725 43860  
 726 43920  
 727 43980  
 728 44040  
 729 44100  
 730 44160  
 731 44220  
 732 44280  
 733 44340  
 734 44400  
 735 44460  
 736 44520  
 737 44580  
 738 44640  
 739 44700  
 740 44760  
 741 44820  
 742 44880  
 743 44940  
 744 45000  
 745 45060  
 746 45120  
 747 45180  
 748 45240  
 749 45300  
 750 45360  
 751 45420  
 752 45480  
 753 45540  
 754 45600  
 755 45660  
 756 45720  
 757 45780  
 758 45840  
 759 45900  
 760 45960  
 761 46020  
 762 46080  
 763 46140  
 764 46200  
 765 46260  
 766 46320  
 767 46380  
 768 46440  
 769 46500  
 770 46560  
 771 46620  
 772 46680  
 773 46740  
 774 46800  
 775 46860  
 776 46920  
 777 46980  
 778 47040  
 779 47100  
 780 47160  
 781 47220  
 782 47280  
 783 47340  
 784 47400  
 785 47460  
 786 47520  
 787 47580  
 788 47640  
 789 47700  
 790 47760  
 791 47820  
 792 47880  
 793 47940  
 794 48000  
 795 48060  
 796 48120  
 797 48180  
 798 48240  
 799 48300  
 800 48360  
 801 48420  
 802 48480  
 803 48540  
 804 48600  
 805 48660  
 806 48720  
 807 48780  
 808 48840  
 809 48900  
 810 48960  
 811 49020  
 812 49080  
 813 49140  
 814 49200  
 815 49260  
 816 49320  
 817 49380  
 818 49440  
 819 49500  
 820 49560  
 821 49620  
 822 49680  
 823 49740  
 824 49800  
 825 49860  
 826 49920  
 827 49980  
 828 50040  
 829 50100  
 830 50160  
 831 50220  
 832 50280  
 833 50340  
 834 50400  
 835 50460  
 836 50520  
 837 50580  
 838 50640  
 839 50700  
 840 50760  
 841 50820  
 842 50880  
 843 50940  
 844 51000  
 845 51060  
 846 51120  
 847 51180  
 848 51240  
 849 51300  
 850 51360  
 851 51420  
 852 51480  
 853 51540  
 854 51600  
 855 51660  
 856 51720  
 857 51780  
 858 51840  
 859 51900  
 860 51960  
 861 52020  
 862 52080  
 863 52140  
 864 52200  
 865 52260  
 866 52320  
 867 52380  
 868 52440  
 869 52500  
 870 52560  
 871 52620  
 872 52680  
 873 52740  
 874 52800  
 875 52860  
 876 52920  
 877 52980  
 878 53040  
 879 53100  
 880 53160  
 881 53220  
 882 53280  
 883 53340  
 884 53400  
 885 53460  
 886 53520  
 887 53580  
 888 53640  
 889 53700  
 890 53760

481

20 Asp Ile Phe Ile Lys Glu Lys Asp Pro Glu Lys Ile Thr Phe Ala Asn.  
35 Thr Thr Gly Leu Ser Arg Thr Phe Lys Val Ser Ala Phe Gln Cys Glu  
50 Gly  
65

(2) INFORMATION FOR SEQ ID NO:329:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 829 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

- (A) NAME/KEY: -  
(B) LOCATION: 1..829  
(D) OTHER INFORMATION: / Ceres Seq. ID 1715962

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:329:

20 aaaaacgcgc tggagaaatc tccctcagg gtctcttga tcaaacaga gacccctcag  
120 ggaagcaaa acaagatcg gagaagatg agacgatct tctctcga acgatgcga  
180 tcccgacag gttaccatc aggttcacg cttaagtgt cgaagtcga agacctcgcg  
240 ggaagctgt tccgatttc aagcatcca actccatt ccagctgatc aggatccag  
300 agctcgaaa gagaagctt agaacatt ct-gggttgg aacagcaaa acaagcctt  
360 ccacgaac cgtctttagc cagctgata actgatct cgtgtacc agaggttcc  
420 gttacaagt gaggttcgt tagccatt tcccatca cgcctccat ggcggtgacg  
480 gaaatctat cgaagctcgt aactcttg ccgagagaa qctgaggaq gtagagatc  
540 tggatcgtt aacatctt cgaactgaga agctaaaga taagattgt ctgacggtc  
600 agacacaga gctgtttca agtctacg ctttctat caagaattt cactcgaaa  
660 agagattat caggaattt ctgatgta tctatgtt cgaqaaac agatcttag  
720 aggaagag aagcacaat tcatcgtt tctcgttt tctgtttt tgaatttag  
780 ttctgttcc agattttt tacttaca caagttttg agacgaagt

(2) INFORMATION FOR SEQ ID NO:330:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 134 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

45 Met Lys Thr Ile Leu Ser Ser Glu Thr Met Asp Ile Pro Asp Ser Val  
1 Thr Ile Lys Val His Ala Lys Val Ile Glu Val Glu Gly Pro Arg Gly  
20 Lys Leu Val Arg Asp Phe Lys His Leu Asn. Leu Asp Phe Gln Leu Ile  
35 Lys Asp Pro Glu Thr Gly Lys Lys Lys Leu Lys Ile Asp Ser Thr Phe  
50 Gly Thr Arg Lys Thr Ser Ala Ser Ile Arg Thr Ala Leu Ser His Val  
65 Asp Asn Leu Ile Ser Gly Val Thr Arg Gly Phe Arg Tyr Lys Met Arg  
80 Phe Val Tyr Ala His Phe Pro Ile Asn Ala Ser Ile Gly Gly Asp Gly  
100 Lys Ser Ile Glu Ile Arg Asn Phe Leu Gly Glu Lys Lys Val Arg Lys  
115 Val Glu Met Leu Asp Gly Val Thr Ile Val Arg Xaa Glu Lys Val Lys

482

130 Asp Glu Ile Val Leu Asp Gly Asn Asp Ile Glu Leu Val Ser Arg Ser  
145 Cys Ala Leu Ile Asn Gln Lys Cys His Val Lys Lys Lys Asp Ile Arg  
160 Lys Phe Leu Asp Gly Ile Tyr Val Ser Glu Lys Ser Lys Ile Val Glu  
175 Glu Glu  
180

10 (2) INFORMATION FOR SEQ ID NO:331:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 185 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

- (A) NAME/KEY: peptide  
(B) LOCATION: 1..185  
(D) OTHER INFORMATION: / Ceres Seq. ID 1715964

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:331:

20 Met Asp Ile Pro Asp Ser Val Thr Ile Lys Val His Ala Lys Val Ile  
1 Met Val Glu Gly Pro Arg Gly Lys Leu Val Arg Asp Phe Lys His Leu  
15 Asn Leu Asp Phe Gln Leu Ile Lys Asp Pro Glu Thr Gly Lys Lys Lys  
30 Leu Lys Ile Asp Ser Thr Phe Gly Thr Arg Lys Thr Ser Ala Ser Ile  
45 Arg Thr Ala Leu Ser His Val Asp Asn Leu Ile Ser Gly Val Thr Arg  
60 Gly Phe Arg Tyr Lys Met Arg Phe Val Tyr Ala His Phe Pro Ile Asn  
75 Ala Ser Ile Gly Gly Asp Gly Lys Ser Ile Glu Ile Arg Asn Phe Leu  
90 Gly Glu Lys Lys Val Arg Lys Val Glu Met Leu Asp Gly Val Thr Ile  
105 Val Arg Xaa Glu Lys Val Lys Asp Glu Ile Val Leu Asp Gly Asn Asp  
120 Ile Glu Leu Val Ser Arg Ser Cys Ala Leu Ile Asn Gln Lys Cys His  
135 Val Lys Lys Lys Asp Ile Arg Lys Phe Leu Asp Gly Ile Tyr Val Ser  
150 Glu Lys Ser Lys Ile Val Glu Glu Glu  
165

(2) INFORMATION FOR SEQ ID NO:332:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 100 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

- (A) NAME/KEY: peptide  
(B) LOCATION: 1..100  
(D) OTHER INFORMATION: / Ceres Seq. ID 1715965

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:332:

55 Met Arg Phe Val Tyr Ala His Phe Pro Ile Asn Ala Ser Ile Gly Gly  
1 Asp Gly Lys Ser Ile Glu Ile Arg Asn Phe Leu Gly Glu Lys Lys Val  
15 Arg Lys Val Glu Met Leu Asp Gly Val Thr Ile Val Arg Xaa Glu Lys

35 Val Lys Asp Glu Ile Val Leu Asp Gly Asn Asp Ile Glu Leu Val Ser 45  
50 Arg Ser Cys Ala Leu Ile Asn Gln Lys Cys His Val Lys Lys Asp 60  
65 Ile Arg Lys Phe Leu Asp Gly Ile Tyr Val Ser Glu Lys Ser Lys Ile 75  
85 Val Glu Glu Glu 90  
100

(2) INFORMATION FOR SEQ ID NO:333:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 675 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

(A) NAME/KEY: -  
(B) LOCATION: 1..675  
(C) OTHER INFORMATION: / Ceres Seq. ID 1508584

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:333:

aactgtcac ctgacagaaa agaaagagc cctagcttt gtcaaaagc ggttcagaaa 60  
caaaaacaa tggcgggaat tggacggata acttagatt gggagccgt ggtgacccgt 120  
aagaataccg ctaacgcgcg tgcacagcgc gacgaagaaa ctgtcaacgc cgtctcaga 180  
tcgggcgcg atacgcgac cgtcagagc tccaatgctg gaacacaaa ggcggatca 240  
agcggacat ctctgaacac aaaaatgctt gatgatgaca ctggaagacc tactatgaa 300  
cgtgtccta ctgacgcaaa gaagccatt atgacagcca gacagacaa gaagctaac 360  
cagctccaa ttctcaaat catcaatgag aagcccaag tgattcaaga gtatgctt 420  
gqcaaaacta taccacaaa gcaatcctt tctaagcttg agagcgctt tggagctag 480  
ctctgggaa agagtggc caagtctac tgatgagca agtaacaga atcaagctt 540  
tgctcaatg ccgaacttt gccaaaga atattctg attgaaga agcaaacccg 600  
ttgaatgt ttgttggtg atggaactc tactcctaa actcatatca atataaac 660  
ttggcttt tcatc

(2) INFORMATION FOR SEQ ID NO:334:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 165 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide  
(B) LOCATION: 1..165  
(C) OTHER INFORMATION: / Ceres Seq. ID 1808585  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:334:  
Asn Leu Ser Pro Ala Glu Lys Glu Arg Ser Pro Arg Phe Cys Gln Lys 1  
Ala Val Ala Glu Gln Lys Thr Met Ala Gly Ile Gly Pro Ile Thr Gln 15  
Asp Trp Glu Pro Val Val Ile Arg Lys Lys Pro Ala Asn Ala Ala Ala 30  
Lys Arg Asp Glu Lys Thr Val Asn Ala Ala Arg Arg Ser Gly Ala Asp 45  
Ile Glu Thr Val Arg Lys Phe Asn Ala Gly Thr Asn Lys Ala Ala Ser 60  
Ser Gly Thr Ser Leu Asn Thr Lys Met Leu Asp Asp Thr Glu Asn 75  
Leu Thr His Glu Arg Val Pro Thr Glu Leu Lys Ala Ile Met Gln 90  
Ala Arg Thr Asp Lys Lys Leu Thr Gln Ser Gln Leu Ala Gln Ile Ile 105  
Asn Glu Lys Pro Gln Val Ile Gln Glu Tyr Glu Ser Gly Lys Ala Ile 120

130 Pro Asn Gln Ile Leu Ser Lys Leu Glu Arg Ala Leu Gly Ala Lys 140  
145 Leu Arg Gly Lys Lys 150  
155

(2) INFORMATION FOR SEQ ID NO:335:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 142 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide  
(B) LOCATION: 1..142  
(C) OTHER INFORMATION: / Ceres Seq. ID 1808586  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:335:  
Met Ala Gly Ile Gly Pro Ile Thr Gln Asp Trp Glu Pro Val Val Ile 1  
Arg Lys Lys Pro Ala Asn Ala Ala Lys Arg Asp Gln Lys Thr Val 15  
Asn Ala Ala Arg Ser Gly Ala Asp Ile Glu Thr Val Arg Lys Phe 30  
Asn Ala Gly Thr Asn Lys Ala Ala Ser Ser Gly Thr Ser Leu Asn Thr 45  
Lys Met Leu Asp Asp Thr Glu Asn Leu Thr His Glu Arg Val Pro 60  
Thr Glu Leu Lys Lys Ala Ile Met Gln Ala Arg Thr Asp Lys Lys Leu 75  
Thr Gln Ser Gln Leu Ala Gln Ile Ile Asn Glu Lys Pro Gln Val Ile 90  
Gln Glu Tyr Glu Ser Gly Lys Ala Ile Pro Asn Gln Cln Ile Leu Ser 105  
Lys Leu Glu Arg Ala Leu Gly Ala Lys Leu Arg Gly Lys Lys 120  
125

(2) INFORMATION FOR SEQ ID NO:336:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 630 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

(A) NAME/KEY: -  
(B) LOCATION: 1..630  
(C) OTHER INFORMATION: / Ceres Seq. ID 1808591  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:336:  
aacactacaa caatttcat accctcatat cttctccatg gaaacacua agcaactctt 60  
ctcttcttc ctatctctg ctatctacac cattatcat ctactctac ttcatattt 120  
cacagaggag caaaacacaa tgcgaacaa aatcatagac gaattggtt caagtgctc 180  
tttgaagat tggagcggag cgtctccaa caacaacac gaattaaay gtccagctc 240  
aacttaact ctctctccc ccaaacatc cgtggaaga atcaacgca ggtcaucat 300  
agttgcttt tactatattg ttccaaatg gcttgacttc tccgtcaa ggtccatgat 360  
gcatttctt cgcattccca cactctcttc tggacactt atagcttta ccaacattc 420  
agctcttgt ttactcttg atgtgttct catctctgag ccagatatt tegtcttcc 480  
tactatglt atccatcgaa tggctttctc attaacttc tctcgtacg gtggtgaga 540  
talaafta ttacttttg attgcttct cactttgtt tcttctact ttactaata 600

(2) INFORMATION FOR SEQ ID NO:337:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 161 amino acids  
(B) TYPE: amino acid

(A) NAME/KEY: -  
(B) LOCATION: 1..630  
(C) OTHER INFORMATION: / Ceres Seq. ID 1808591  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:337:

aacactacaa caatttcat accctcatat cttctccatg gaaacacua agcaactctt 60  
ctcttcttc ctatctctg ctatctacac cattatcat ctactctac ttcatattt 120  
cacagaggag caaaacacaa tgcgaacaa aatcatagac gaattggtt caagtgctc 180  
tttgaagat tggagcggag cgtctccaa caacaacac gaattaaay gtccagctc 240  
aacttaact ctctctccc ccaaacatc cgtggaaga atcaacgca ggtcaucat 300  
agttgcttt tactatattg ttccaaatg gcttgacttc tccgtcaa ggtccatgat 360  
gcatttctt cgcattccca cactctcttc tggacactt atagcttta ccaacattc 420  
agctcttgt ttactcttg atgtgttct catctctgag ccagatatt tegtcttcc 480  
tactatglt atccatcgaa tggctttctc attaacttc tctcgtacg gtggtgaga 540  
talaafta ttacttttg attgcttct cactttgtt tcttctact ttactaata 600

(2) INFORMATION FOR SEQ ID NO:338:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 161 amino acids  
(B) TYPE: amino acid

485

(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(11) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..181  
(D) OTHER INFORMATION: / Ceres Seq. ID 1808592  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:337:  
Thr Leu Gln Gln Phe Ser Tyr Leu Ile Phe Ser Ser Met Ala Thr Thr  
1 10 15  
Lys His Leu Leu Phe Phe Leu Ile Ile Ala Ile Ile Thr Ile Thr  
20 25 30  
Ser Ser Thr Ser Leu Pro Phe Leu Thr Thr Glu Gln Asn Gln Ile Ala  
35 40 45  
Thr Lys Ile Ile Asp Ala Met Val Ser Ser Gly Ser Phe Glu Asp Trp  
50 55 60  
Ser Gly Ala Phe Leu Asn Asn Asp Glu Leu Asn Gly Pro Val Leu  
65 70 75 80  
Thr Ser Thr Leu Phe Leu Pro Lys Thr Ser Val Glu Gly Ile Asn Ala  
85 90 95  
Thr Ser Pro Leu Val Ala Ser Tyr His Ile Val Pro Gln Trp Leu Asp  
100 105 110  
Phe Ser Val Ile Ser Leu Met Met Pro Phe Ser Arg Ile Pro Thr Leu  
115 120 125  
Leu Ser Gly His Ser Ile Val Val Thr Asn Asn Ser Ala Ser Gly Phe  
130 135 140  
Thr Leu Asp Gly Val Leu Ile Ser Glu Pro Asp Leu Phe Val Ser Pro  
145 150 155 160  
Thr Ile Val Ile His Arg Met Ala Phe Pro Phe Asn Phe Ser Arg Tyr  
165 170 175  
Gly Gly Asp Ile  
180

(2) INFORMATION FOR SEQ ID NO:338:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 169 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: Peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..169

(D) OTHER INFORMATION: / Ceres Seq. ID 1808593

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:338:

Met Ala Thr Thr Lys His Leu Leu Phe Leu Leu Leu Ile Ile Ala Ile  
1 5 10 15  
Ile Thr Ile Thr Ser Ser Thr Ser Leu Pro Phe Leu Thr Thr Glu Gln  
20 25 30  
Asn Gln Ile Ala Thr Lys Ile Ile Asp Ala Met Val Ser Ser Gly Ser  
35 40 45  
Phe Glu Asp Trp Ser Gly Ala Phe Leu Asn Asn Asp Glu Leu Asn  
50 55 60  
Gly Pro Val Leu Thr Ser Thr Leu Phe Leu Pro Lys Thr Ser Val Glu  
65 70 75 80  
Gly Ile Asn Ala Thr Ser Pro Leu Val Ala Ser Tyr His Ile Val Pro  
85 90 95  
Gln Trp Leu Asp Phe Ser Val Ile Ser Leu Met Met Pro Phe Ser Arg  
100 105 110  
Ile Pro Thr Leu Leu Ser Gly His Ser Ile Val Val Thr Asn Asn Ser  
115 120 125  
Ala Ser Gly Phe Thr Leu Asp Gly Val Leu Ile Ser Glu Pro Asp Leu  
130 135 140

486

Phe Val Ser Pro Thr Ile Val Ile His Arg Met Ala Phe Pro Phe Asn  
145 150 155 160  
Phe Ser Arg Tyr Gly Gly Asp Ile  
165  
(2) INFORMATION FOR SEQ ID NO:339:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 127 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(11) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..127  
(D) OTHER INFORMATION: / Ceres Seq. ID 1808594  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:339:  
Met Val Ser Ser Gly Ser Phe Glu Asp Trp Ser Gly Ala Phe Leu Asn  
1 5 10 15  
Asn Asn Asp Glu Leu Asn Gly Pro Val Leu Thr Ser Thr Leu Phe Leu  
20 25 30  
Pro Lys Thr Ser Val Glu Gly Ile Asn Ala Thr Ser Pro Leu Val Ala  
35 40 45  
Ser Tyr His Ile Val Pro Gln Trp Leu Asp Phe Ser Val Ile Ser Leu  
50 55 60  
Met Met Pro Phe Ser Arg Ile Pro Thr Leu Leu Ser Gly His Ser Ile  
65 70 75 80  
Val Val Thr Asn Asn Ser Ala Ser Gly Phe Thr Leu Asp Gly Val Leu  
85 90 95  
Ile Ser Glu Pro Asp Leu Phe Val Ser Pro Thr Ile Val Ile His Arg  
100 105 110 115  
Met Ala Phe Pro Phe Asn Phe Ser Arg Tyr Gly Gly Asp Ile  
120 125  
(2) INFORMATION FOR SEQ ID NO:340:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 717 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(11) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..717  
(D) OTHER INFORMATION: / Ceres Seq. ID 1920563  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:340:  
ctgtttata ccaatgacca agagagact gtaagagg ttgtccaaa ttctttgtgt  
60  
gtttggcac ccgaagtgga tgttaagctg actggtctcg taggcaaga atgtcttctg  
120  
ccaagatc caaacgccac cgttaattatg ctgtccnag ggaaggaat tgcctcttc  
180  
aggtcttct tatgaagat gttctttgag aaactgatg actacaagt caatgctta  
240  
gcttgggtgt tcttgggtgt accaacact agctcaatgc tctaccaga ggaattgat  
300  
agagaaga caaagcccc cgaacttc aggttggtt accgataag ccgagaacaa  
360  
gcgaacata agggagagaa atgcatatc cagactcggg tagcacagta ccgacgta  
420  
ttatggagt tgttaagaa agacaact ttgtttaca tgtgtggac caagggaatg  
480  
gagaagaa ttgatgacat tatgttcca ttgctgcaa atgacggatc tgaatggtt  
540  
gattacaaga agcagttgaa gaagccagag caatggacq ttgaatcta ctatcaaaa  
600  
agctttgac attctctag caaatatag ctgaacaaa ctgtaattt cgttccgaa  
660  
ttctgtatt ttgaadtaa gttttttaga tatgttatac taataaaga gttcttt  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 197 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

487

(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..197  
(D) OTHER INFORMATION: / Ceres Seq. ID 1920564  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:341:  
Leu Val Tyr Thr Asn Asp Gln Gly Glu Thr Val Lys Gly Val Cys Ser  
1 5 10 15  
Asn Phe Leu Cys Asp Leu Ala Pro Gly Ser Asp Val Lys Leu Thr Gly  
20 25 30  
Pro Val Gly Lys Glu Met Leu Met Pro Lys Asp Pro Asn Ala Thr Val  
35 40 45  
Ile Met Leu Ala Thr Gly Thr Gly Ile Ala Pro Phe Arg Ser Phe Leu  
50 55 60  
Trp Lys Met Phe Phe Glu Lys His Asp Asp Tyr Lys Phe Asn Gly Leu  
65 70 75  
Ala Trp Leu Phe Leu Gly Val: Pro Thr Thr Ser Ser Leu Leu Tyr Gln  
80 85 90  
Glu Glu Phe Asp Lys Met Lys Ala Lys Ala Pro Glu Asn Phe Arg Val  
95 100  
Asp Tyr Ala Ile Ser Arg Glu Ala Asn Asp Lys Gly Glu Lys Met  
105 110 115 120  
Tyr Ile Gln Thr Arg Met Ala Gln Tyr Ala Ala Glu Leu Trp Glu Leu  
125 130 135 140  
Leu Lys Lys Asp Asn Thr Phe Val Tyr Met Cys Gly Leu Lys Gly Met  
145 150 155  
Glu Lys Gly Ile Asp Asp Ile Met Val Ser Leu Ala Ala Asn Asp Gly  
160 165 170 175  
Ile Asp Trp Phe Asp Tyr Lys Lys Gln Leu Lys Lys Ala Glu Gln Trp  
180 185 190  
Asn Val Glu Val Tyr  
195

(2) INFORMATION FOR SEQ ID NO:342:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 160 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..160  
(D) OTHER INFORMATION: / Ceres Seq. ID 1920565  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:342:  
Met Leu Met Pro Lys Asp Pro Asn Ala Thr Val Ile Met Leu Ala Thr  
5 10 15  
Gly Thr Gly Ile Ala Pro Phe Arg Ser Phe Leu Trp Lys Met Phe Phe  
20 25 30  
Glu Lys His Asp Asp: Tyr Lys Phe Asn Gly Leu Ala Trp Leu Phe Leu  
35 40 45  
Gly Val Pro Thr: Thr Ser Ser: Leu Leu Tyr Gln Glu Gly Phe Asp Lys  
50 55 60  
Met Lys Ala Lys Ala Pro Glu Asn Phe Arg Val Asp Tyr Ala Ile Ser  
65 70 75 80  
Arg Glu Gln Ala Asn Asp Lys Gly Glu Lys Met Tyr Ile Gln Thr Arg  
85 90 95  
Met Ala Gln Tyr Ala Ala Glu Leu Trp Glu Leu Lys Lys Asp Asn  
100 105 110  
Thr Phe Val Tyr Met Cys Gly Leu Lys Gly Met Glu Lys Gly Ile Asp  
115 120 125  
Asp Ile Met Val Ser Leu Ala Ala Asn Asp Gly Ile Asp Trp Phe Asp  
130 135 140

488

Tyr Lys Lys Gln Leu Lys Lys Ala Glu Gln Trp Asn Val Glu Val Tyr  
145 150 155 160

(2) INFORMATION FOR SEQ ID NO:343:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 158 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..158  
(D) OTHER INFORMATION: / Ceres Seq. ID 1920566  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:343:  
Met Pro Lys Asp Pro Asn Ala Thr Val Ile Met Leu Ala Thr Gly Thr  
1 5 10 15  
Gly Ile Ala Pro Phe Arg Ser Phe Leu Trp Lys Met Phe Phe Glu Lys  
20 25 30  
His Asp Asp Tyr Lys Phe Asn Gly Leu Ala Trp Leu Phe Leu Gly Val  
35 40 45  
Pro Thr Thr Ser Ser Leu Leu Tyr Gln Glu Glu Phe Asp Lys Met Lys  
50 55 60  
Ala Lys Ala Pro Glu Asn Phe Arg Val Asp Tyr Ala Ile Ser Arg Glu  
65 70 75 80  
Gln Ala Asn Asp Lys Gly Glu Lys Met Tyr Ile Gln Thr Arg Met Ala  
85 90 95  
Gln Tyr Ala Ala Glu Leu Trp Glu Leu Lys Lys Asp Asn Thr Phe  
100 105 110  
Val Tyr Met Cys Gly Leu Lys Gly Met Glu Lys Gly Ile Asp Asp Ile  
115 120 125  
Met Val Ser Leu Ala Ala Asn Asp Gly Ile Asp Trp Phe Asp Tyr Lys  
130 135 140 145  
Lys Gln Leu Lys Lys Ala Glu Gln Trp Asn Val Glu Val Tyr  
150 155

(2) INFORMATION FOR SEQ ID NO:344:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2192 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..2192  
(D) OTHER INFORMATION: / Ceres Seq. ID 1974419  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:344:  
aattcaagtc atctctatcc gtcaacaata caaaccaacc tctctaatc ctctctttc 60  
atccacattt tctactagta atgcttcta gtcaatgag cactcatca tggacagcta 120  
gagagacaa gcaattcga atgctgttqg cgaattcca caggacact ctgacggtt 180  
ggcaaaaaa tgcaggaga gtctgtgga aatcaacta agaagtaag cctacattat 240  
aattgtctct taggaatgtg aatgacattg agtcaggag ctatccaaa cctgattat 300  
gcttgctgc ggcattagt gtgctgttca ctcaatgtc gattgcacg aacgcgctt 360  
acgtgactc cgaatttgc ttctcttct tctcttctt accctcgca gaagaatctc 420  
caacgatca taattttct tgaatttta aatcgagac taacctgac tctgtgagc 480  
tcaatgagtc tggttctgat cctgagtcct tggaaaggag tgcataagct ctctgaaa 540  
ctaatgactc tctctatcc aacaggtgt ttgatctaat gcgaagcag gagaagactc 600  
gattagtga attagcgtt gagaagaac ataatgaac tattcaagc atgaagaca 660  
tgaagaca cgcgaattg gcaggagtc agagaattt agtcagcaa caggcccaag 720  
cgaagctca aaattctaga tatggagtg agtggccag gaaggaatg cagacagata 780  
atgaagctca gacagagcat aatgttgat tggttctgat gcgaagcag tctctatc 840  
ggaagagaa agcaagaatt gccacggag acacagatca agcaagcag cgcgagctg 900  
agaagagag agcgaactt gacgagaaa caattcgct gaaagctatg gctgagcgtg 960

489  
aagccgacg tcatgaagct aactcactg agacacagaa tagaagaatg cttctagata 1020  
agataaagg tgaaggagg aatggcttg cagcaatcaa cagcacttc agtcacattg 1080  
aaggaggag- aaggactta ttaactgac gaagcaaat gattatgact gttggaggag 1140  
ttacggctg- agccctggg attacaaa ctgagaagg tttagaggtt acctggggtt 1200  
atacaatag aatctctgg cagcacaac tcaracaga atctctcatg ggcgcagctc 1260  
catggcgag ctaagtgtt cagtttaaga caaactctg cacagctga gggcgagcag 1320  
catctcgaga aggaagaag cctctgaaa atg-shuttct cctcgtctt ttgaagcaga 1380  
gaattagctg ttctgcaag gccacagcaa ataccagtc acataagcaa ccatcccgaa 1440  
acatagct- ttaaggact caggcgaccg gaaaactat gttggcgagg gagattgctc 1500  
ggaagtcggg tct-gaattat gcta-gatga caggagagaga tgttgacct cttgctgcan 1560  
aggtctttac aaagatccat gaagatttg atgggtctaa gaaatcaaac aaagggttac 1620  
tgtctttac- cga-gaagct gatgcttcc tatcggaagc taacagact taactagtg 1680  
aggtcagctg cagcgtctg aacgcttgc tctttcgaa cgggtgatcaa tctcgagaca 1740  
tagctctg- cctggctaca aacagactg gagatctga cagtgagtc actgcagaga 1800  
tcgacgaagt tarcgagtt cctcctctg tgaagaaga acgtctcaag cctctcaagg 1860  
tctatctcaa caagctacta atgggtgagc acagaagaag tgaagaagac taaacctta 1920  
aatggaccaa ctgttcaag aagaagaat cacaagaagt aaccattgaa ggaaccttaa 1980  
cgaacaaag- gat-aaaga actgcnaaaa agacagaag cttctcgtt cgtgaatcg 2040  
ccaaacttg- tggggagtt cagctcgagg tatatggagc acagatattt gttcggatt 2100  
tgcgaactga agt-ggcca tegtctcgt ag

(1) INFORMATION FOR SEQ ID NO:345:  
(1) SEQUENCE CHARACTERISTICS:  
(i) LENGTH: 703 amino acids  
(ii) TYPE: amino acid  
(iii) STRANDEDNESS:  
(iv) TOPOLOGY: linear  
(11) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide  
(B) LOCATION: 1..703  
(D) OTHER INFORMATION: / Ceres Seq. ID 1974420

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:345:  
Met Ala Ser Ser Ser Met Ser Thr Ser Ser Thr Trp Thr Ala Arg Glu Asp  
Lys Gln Phe Glu Met Ala Leu Ala Lys Phe Asp Lys Asp Thr Pro Asp  
Arg Trp Gln Lys Ile Ala Arg Ala Val Gly Gly Lys Ser Thr Glu Glu  
Val Lys Arg His Tyr Glu Leu Leu Arg Asp Val Asn Asp Ile Glu  
Ser Gly Arg Tyr Pro Gln Pro Arg Leu Cys Ser Ala Ala Ile Ala  
Ala Ala Phe Thr Ser Met Ser Met Ser Gl: Asn Arg Ala Tyr Ala Asp  
Ser Arg Phe Arg Phe Pro Phe Ser Ser Ser Pro Pro Ala Glu Glu  
Ser Pro Thr Arg His Lys Ser Ser Ser Asn: Ser Lys Ser Glu Thr Lys  
Pro Asp Ser Asp Glu Pro Lys Gly Ser Gly Phe Asp Pro Glu Ser Leu  
Glu Arg Gly Ala Lys Ala Leu Arg Glu Ile Asn Ser Ser Pro His Ser  
Lys Gln Val Phe Asp Leu Met Arg Lys Gl: Glu Lys Thr Arg Leu Ala  
Glu Leu Ala Ala Glu Lys Glu His Asn Gl: Ala Ile Gln Ala Ser Lys  
Asp Ile Glu Arg Gln Arg Lys Leu Ala Glu Asp Gln Arg Asn Leu Val  
Gln Gln Gln Ala Gln Ala Lys Ala Gln Asn: Leu Arg Tyr Glu Asp Glu  
Leu Ala Arg Lys Arg Met Gln Thr Asp Asn Glu Ala Gln Arg Arg His

225 230 235 490 240  
Asn Ala Glu Leu Val Ser Met Gln Glu Ala Ser Ile Arg Lys Glu  
Lys Ala Arg Ile Ala Thr Glu Glu Gln Ile Gln Ala Gln Gln Arg Glu  
Thr Glu Lys Glu Arg Ala Glu Leu Glu Thr Ile Arg Val Lys  
Ala Met Ala Glu Ala Glu Gly Arg Ala His Glu Ala Lys Leu Thr Glu  
Glu Gln Asn Arg Arg Met Leu Leu Asp Lys Ile Asn Gly Glu Arg Glu  
Lys Trp Leu Ala Ala Ile Asn Thr Thr Phe Ser His Ile Glu Gly Glu  
Val Arg Thr Leu Leu Thr Asp Arg Ser Lys Leu Ile Met Thr Val Gly  
Gly Val Thr Ala Leu Ala Ala Gly Val Tyr Thr Arg Glu Gly Ala  
Arg Val Thr Trp Gly Tyr Ile Asn Arg Ile Leu Gly Gln Pro Ser Leu  
Ile Arg Glu Ser Ser Met Gly Arg Phe Pro Trp Ala Gly Ser Val Ser  
Gln Phe Lys Asn Lys Leu Ser: Thr Ala Ala Gly Ala Ala Ser Ala  
Glu Gly Glu Lys Pro Leu Glu Asn Val Ile Leu His Arg Ser Leu Lys  
Thr Arg Ile Glu Arg Leu Ala Arg Ala Thr Ala Asn Thr Lys Ser: His  
Lys Ala Pro Phe Arg Asn Met: Met Phe Tyr Gly Pro Gly Thr Gly  
Lys Thr Met Val Ala Arg Glu Ile Ala Arg Lys Ser Gly Leu Asp Tyr  
Ala Met Met Thr Gly Gly Asp Val Ala Pro Leu Leu Ala Gln Ala Val  
Thr Lys Ile His Glu Ile Phe Asp Trp Ala Lys Lys Ser Asn Lys Gly  
Leu Leu Leu Phe Ile Asp Glu Ala Asp Ala Phe Leu Cys Glu Arg Asn  
Ser Thr Tyr: Met Ser Glu Ala Gln Arg Ser Ala Leu Asn Ala Leu Leu  
Phe Arg Thr Gly Asp Gln Ser Arg Asp Ile Val Leu Val Leu Ala Thr  
Asn Arg Pro Gly Asp Leu Ser Ala Val Thr Asp Arg Ile Asp Glu  
Val Ile Glu Phe Pro Leu Pro Gly Glu Glu Arg Phe Lys Leu Leu  
Lys Leu Tyr Leu Asn Lys Tyr Leu Met Gly Asp Asp Lys Lys Gly Glu  
Lys Asp Ser Asn Leu Lys Trp Ser Asn Leu Phe Lys Lys Lys Ser  
Gln Lys Ile Thr Ile Glu Gly Asp Leu Thr Asp Gln Val Ile Lys Glu  
Ala Ala Lys Lys Thr Glu Gly Phe Ser Gly Arg Glu Ile Ala Lys Leu  
Val Ala Gly Val: Gln Ala Ala Val Tyr Gly Arg Gln Asp Cys Val Leu  
Asp Ser Gln Leu Phe Glu Glu Ile Val Asp Tyr Lys Ile Glu Glu His  
His Gln Arg Ile Arg Leu Ala Thr Glu Gly Gln Ser: Phe Pro

(2) INFORMATION FOR SEQ ID NO:346:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 698 amino acids



491

(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..698  
(D) OTHER INFORMATION: / Ceres Seq. ID 1974421  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:346:  
1 Met Ser Thr Ser Ser Trp Thr Ala Arg Glu Asp Lys Gln Phe Glu Met 15  
Ala Leu Ala Lys Phe Asp Lys Asp Thr Pro Asp Arg Trp Gln Lys Ile 30  
Ala Arg Ala Val Gly Gly Lys Ser Thr Glu Gln Val Lys Arg His Tyr 45  
Glu Leu Leu Leu Arg Asp Val Asn Asp Ile Gln Ser Gly Arg Tyr Pro 60  
Gln Pro Arg Leu Cys Ser Ala Ala Ala Ile Ala Ala Phe Thr Ser 75  
Met Ser Met Ser Gln Asn Arg Ala Tyr Ala Asp Ser Arg Phe Arg Phe 90  
Pro Phe Phe Ser Ser Pro Pro Ala Glu Gln Ser Pro Thr Asp His 105  
Lys Ser Ser Asn Ser Lys Ser Glu Thr Lys Pro Asp Ser Asp Glu 120  
Pro Lys Gly Ser Gly Phe Asp Pro Glu Ser Leu Glu Arg Gly Ala Lys 135  
Ala Leu Arg Glu Ile Asn Ser Ser Pro His Ser Lys Gln Val Phe Asp 150  
Leu Met Arg Lys Gln Glu Lys Thr Arg Leu Ala Glu Leu Ala Ala Glu 165  
Lys Glu His Asn Gln Ala Ile Gln Ala Ser Lys Asp Ile Glu Arg Gln 180  
Arg Lys Leu Ala Glu Asp Gln Arg Asn Leu Val Gln Gln Gln Ala Gln 195  
Ala Lys Ala Gln Asn Leu Arg Tyr Glu Asp Gln Leu Ala Arg Lys Arg 210  
Met Gln Thr Asp Asn Glu Ala Gln Arg Arg His Asn Ala Glu Leu Val 225  
Ser Met Gln Glu Ala Ser Ile Arg Lys Gln Lys Ala Arg Ile Ala 240  
Thr Glu Glu Gln Ile Gln Ala Gln Arg Gln Thr Glu Lys Glu Arg 255  
Ala Glu Leu Glu Arg Glu Thr Ile Arg Val Lys Ala Met Ala Glu Ala 270  
Glu Gly Arg Ala His Glu Ala Lys Leu Thr Glu Glu Gln Asn Arg Arg 285  
Met Leu Leu Asp Lys Ile Asn Gly Glu Arg Glu Lys Trp Leu Ala Ala 300  
Ile Asn Thr Thr Phe Ser His Ile Glu Gly Gly Val Arg Thr Leu Leu 315  
Thr Asp Arg Ser Lys Leu Ile Met Thr Val Gly Gly Val Thr Ala Leu 330  
Ala Ala Gly Val Tyr Thr Arg Glu Gly Ala Arg Val Thr Trp Gly 345  
Tyr Ile Asn Arg Ile Leu Gly Gln Pro Ser Leu Ile Arg Glu Ser Ser 360  
Met Gly Arg Phe Pro Trp Ala Gly Ser Val Ser Gln Phe Lys Asn Lys 375  
385 Ser Thr Ala Ala Ala Ala Ser Ala Glu Gly Glu Lys Pro 400  
Leu Glu Asn Val Ile Leu His Arg Ser Leu Lys Thr Arg Ile Glu Arg 415

492

420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495  
Leu Ala Arg Ala Thr Ala Asn Thr Lys Ser His Lys Ala Pro Phe Arg 420  
Asn Met Met Phe Tyr Gly Pro Gly Thr Gly Lys Thr Met Val Ala 435  
Arg Glu Ile Ala Arg Lys Ser Gly Leu Asp Tyr Ala Met Met Thr Gly 450  
Gly Asp Val Ala Pro Leu Gly Ala Gln Ala Val Thr Lys Ile His Glu 465  
Ile Phe Asp Trp Ala Lys Lys Ser Asn Lys Gly Leu Leu Phe Ile 480  
Asp Glu Ala Asp Ala Phe Leu Cys Glu Arg Asn Ser Thr Tyr Met Ser 495  
Glu Ala Gln Arg Ser Ala Leu Asn Ala Leu Leu Phe Arg Thr Gly Asp 510  
Gln Ser Arg Asp Ile Val Leu Val Leu Ala Thr Asn Arg Pro Gly Asp 525  
Leu Asp Ser Ala Val Thr Asp Arg Ile Asp Glu Val Ile Glu Phe Pro 540  
Leu Pro Gly Glu Glu Arg Phe Lys Leu Lys Leu Lys Leu Tyr Leu Asn 555  
Lys Tyr Leu Met Gly Asp Asp Lys Lys Gly Glu Lys Asp Ser Asn Leu 570  
Lys Trp Ser Asn Leu Phe Lys Lys Lys Ser Gln Lys Ile Thr Ile 585  
Glu Gly Asp Leu Thr Asp Gln Val Ile Lys Glu Ala Ala Lys Lys Thr 600  
Glu Gly Phe Ser Gly Arg Gln Ile Ala Lys Leu Val Ala Gly Val Gln 615  
Ala Ala Val Tyr Gly Arg Gln Asp Cys Val Leu Asp Ser Gln Leu Phe 630  
Glu Glu Ile Val Asp Tyr Lys Ile Glu Glu His His Gln Arg Ile Arg 645  
Leu Ala Thr Glu Gly Gln Ser Phe Pro 660  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 683 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..683  
(D) OTHER INFORMATION: / Ceres Seq. ID 1974422  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:347:  
Met Ala Leu Ala Lys Phe Asp Lys Asp Thr Pro Asp Arg Trp Gln Lys 1  
Ile Ala Arg Ala Val Gly Gly Lys Ser Thr Glu Glu Val Lys Arg His 15  
Tyr Glu Leu Leu Leu Arg Asp Val Asn Asp Ile Glu Ser Gly Arg Tyr 30  
Pro Gln Pro Arg Leu Cys Ser Ala Ala Ile Ala Ala Phe Thr 45  
Ser Met Ser Met Ser Gln Asn Arg Ala Tyr Ala Asp Ser Arg Phe Arg 60  
Phe Pro Phe Phe Ser Ser Pro Pro Ala Glu Glu Ser Pro Thr Asp 75  
His Lys Ser Ser Asn Ser Lys Ser Glu Thr Lys Pro Asp Ser Asp 90  
Glu Pro Lys Gly Ser Gly Phe Asp Pro Glu Ser Leu Glu Arg Gly Ala 105

493

115 Lys Ala Leu Arg Glu Ile Asn Ser Ser Pro His Ser Lys Gln Val Phe 125  
 133 135 140  
 5 Asp Leu Met Arg Lys Gln Glu Lys Thr Arg Leu Ala Glu Leu Ala Ala 150  
 145 Glu Lys Glu His Asn Glu Ala Ile Gln Ala Ser Lys Asp Ile Glu Arg 160  
 165 170 175  
 Gln Arg Lys Leu Ala Glu Asp Gln Arg Asn Leu Val Gln Gln Gln Ala 180  
 180 185 190  
 10 Gln Ala Lys Ala Gln Asn Leu Arg Tyr Glu Asp Glu Leu Ala Arg Lys 195  
 195 200 205  
 Arg Met Gln Thr Asp Asn Glu Ala Gln Arg Arg His Asn Ala Glu Leu 210  
 210 215 220  
 15 Val Ser Met Gln Glu Ala Ser Ile Arg Lys Glu Lys Ala Arg Ile 225  
 225 230 235  
 Ala Thr Glu Glu Gln Ile Gln Ala Gln Gln Arg Glu Thr Glu Lys Glu 240  
 245 250 255  
 Arg Ala Glu Leu Glu Arg Glu Thr Ile Arg Val Lys Ala Met Ala Glu 260  
 260 265 270  
 20 Ala Glu Gly Arg Ala His Glu Ala Lys Leu Thr Glu Glu Gln Asn Arg 275  
 275 280 285  
 Arg Met Leu Leu Asp Lys Ile Asn Gly Glu Arg Glu Lys Trp Leu Ala 290  
 290 295 300  
 25 Ala Ile Asn Thr Thr Phe Ser His Ile Glu Gly Gly Val Arg Thr Leu 305  
 310 315 320  
 Leu Thr Asp Arg Ser Lys Leu Ile Met Thr Val Gly Gly Val Thr Ala 325  
 325 330 335  
 Leu Ala Ala Gly Val Tyr Thr Thr Arg Glu Gly Ala Arg Val Thr Trp 340  
 340 345 350  
 30 Gly Tyr Ile Asn Arg Ile Leu Gly Gln Pro Ser Leu Ile Arg Glu Ser 355  
 355 360 365  
 Ser Met Gly Arg Phe Pro Trp Ala Gly Ser Val Ser Gln Phe Lys Asn 370  
 370 375 380  
 35 Lys Leu Ser Thr Ala Ala Gly Ala Ala Ser Ala Glu Gly Glu Lys 385  
 390 395 400  
 Pro Leu Glu Asn Val Ile Leu His Arg Ser Leu Lys Thr Arg Ile Glu 405  
 410 415  
 Arg Leu Ala Arg Ala Thr Ala Asn Thr Lys Ser His Lys Ala Pro Phe 420  
 420 425 430  
 40 Arg Asn Met Met Phe Tyr Gly Pro Gly Thr Gly Lys Thr Met Val 435  
 435 440 445  
 Ala Arg Glu Ile Ala Arg Lys Ser Gly Leu Asp Tyr Ala Met Thr 450  
 450 455 460  
 45 Gly Gly Asp Val Ala Pro Leu Gly Ala Gln Ala Val Thr Lys Ile His 465  
 470 475 480  
 Glu Ile Phe Asp Trp Ala Lys Lys Ser Asn Lys Gly Leu Leu Phe 485  
 490 495  
 Ile Asp Glu Ala Asp Ala Phe Leu Cys Gln Arg Asn Ser Thr Tyr Met 500  
 500 505 510  
 50 Ser Glu Ala Gln Arg Ser Ala Leu Asn Ala Leu Leu Phe Arg Thr Gly 515  
 520 525  
 Asp Gln Ser Arg Asp Ile Val Leu Val Leu Ala Thr Asn Arg Pro Gly 530  
 530 535 540  
 55 Asp Leu Asp Ser Ala Val Thr Asp Arg Ile Asp Glu Val Ile Glu Phe 545  
 550 555 560  
 Pro Leu Pro Gly Glu Glu Arg Phe Lys Leu Leu Lys Leu Tyr Leu 565  
 570 575  
 Asn Lys Tyr Leu Met Gly Asp Asp Lys Lys Gly Glu Lys Asp Ser Asn 580  
 580 585 590  
 60 Leu Lys Trp Ser Asn Leu Phe Lys Lys Lys Ser Gln Lys Ile Thr 595  
 600 605  
 Ile Glu Gly Asp Leu Thr Asp Gln Val Ile Lys Glu Ala Ala Lys Lys

494

610 615 620  
 Thr Glu Gly Phe Ser Gly Arg Glu Ile Ala Lys Leu Val Ala Gly Val 625  
 630 635 640  
 5 Gln Ala Ala Val Tyr Gly Arg Gln Asp Cys Val Leu Asp Ser Gln Leu 645  
 650 655  
 Phe Glu Glu Ile Val Asp Tyr Lys Ile Glu His His Gln Arg Ile 660  
 665 670  
 Arg Leu Ala Thr Glu Gly Gln Ser Phe Pro 675  
 680 685  
 10 (2) INFORMATION FOR SEQ ID NO:348:  
 (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 953 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (11) MOLECULE TYPE: DNA (genomic)  
 (1x) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1...953  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1975983  
 (x1) SEQUENCE DESCRIPTION: SEQ ID NO:348:  
 naaaagata aacacaaagc tgaiaagatg aaacacagag aaggaacaga caacacacac  
 caacacagat ctacagctctt ccttaaccca tgaacacaaa ggcgcacccg ttacccaagt  
 ttgaagagac tggatgacgt tgggtggtg agtcgcctc ctcttcaat ctacacgaa  
 tcatgcat atccagagat aatgcgcgtg agatctccaa gacaccttc aaagaggtg  
 ttgtgttg caaaaagat tcaatctac caacacccc ttgat-gag agtgttcaa  
 atccgaagt tacaaggtg atgcagagtt tcaatctaa ggaatagtg agagagacct  
 ttgttggt gcatctac tggttccca caaatgaag ttatgacct tttagacct  
 acctaatct ccatctag atgttctgc cctctctaa gaagcaacag aagctcttg  
 gtcacatt tgaagtggt gatgacgct ccagtgcac tctctggt gatgagaga  
 ggaagtgg tgacagagt gataccgtg gaggctctaa atcagtgga gatgagtg  
 acaaggtgg agcaatgct gtatccacc ctgcctcag ggtgagat agtgagaaa  
 ggcagaggt tgggtgga gctggtggt ttggtggtg tctgttcca gctctggt  
 tgcatacacc ttgaagaag cttctgtgt tcttctggt ctatttaag gtacatagc  
 acctattga gaacagagt gtcattgga acatgcttc ttctcttaa accttttac  
 aatgatgt tttttttag ttgaattt atgaagaa gtgtgagctg tacaatgac  
 tttacact gctcctct ttaaatctc ctatttautl tttttcagag ctt  
 60  
 120  
 180  
 240  
 300  
 360  
 420  
 480  
 540  
 600  
 660  
 720  
 780  
 840  
 900

(1) INFORMATION FOR SEQ ID NO:349:  
 (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 214 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (11) MOLECULE TYPE: peptide  
 (1x) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1...214  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1975984  
 (x1) SEQUENCE DESCRIPTION: SEQ ID NO:349:  
 Met Ala Asp Lys Gly Arg Pro Leu Pro Lys Phe Gly Arg Ala Glu Met  
 1 5 10 15  
 Arg Cys Gly Gly Val Ala Ser Phe Asn Leu Ile Ala Ile Met  
 20 25 30  
 Ile Ile Ser Glu Thr Asn Arg Arg Glu Ile Ser Lys Tyr Leu Phe Lys  
 35 40 45  
 Glu Gly Val Leu Phe Ala Lys Lys Asp Phe Asn Leu Pro Gln His Pro  
 50 55 60  
 Leu Ile Glu Ser Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln Ser  
 65 70 75 80  
 Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His Tyr  
 85 90 95  
 Tyr Trp Phe Leu Thr Asn Glu Gly Ile Asp Phe Leu Arg Thr Tyr Leu

```

495
100 105 110
Asn Leu Pro Ser Glu Ile Val Pro Ala Thr Leu Lys Lys Gln Gln Lys
115 120 125
Pro Leu Gly Arg Pro Phe Gly Gly Gly Gly Asp Arg Pro Arg Gly Pro
5 130 135 140
Pro Arg Gly Asp Gly Glu Arg Arg Phe Gly Asp Arg Asp Gly Tyr Arg
145 150 155 160
Gly Gly Pro Lys Ser Gly Gly Glu Tyr Gly Asp Lys Ala Gly Ala Pro
165 170 175
10 Ala Asp Tyr Gln Pro Gly Phe Arg Gly Gly Ala Ser Gly Ala Arg Gln
180 185 190
Gly Phe Gly Arg Gly Ala Gly Gly Phe Gly Gly Gly Ala Gly Pro Ala
195 200 205
Ala Gly Ser Asp Leu Pro
15 210
(2) INFORMATION FOR SEQ ID NO:350:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 199 amino acids
(B) TYPE: amino acid
20 (C) STRANDEDNESS:
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: peptide
(ix) FEATURE:
(A) NAME/KEY: peptide
25 (B) LOCATION: 1..199
(D) OTHER INFORMATION: / Ceres Seq. ID 1975985
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:350:
Met Arg Cys Gly Gly Gly Val Ala Ser Phe Phe Asn Leu Ile Ala Ile
1 5 10 15
30 Met Ile Ile Ser Glu Thr Asn Arg Arg Glu Ile Ser Lys Tyr Leu Phe
20 25 30
Lys Glu Gly Val Leu Phe Ala Lys Lys Asp Phe Asn Leu Pro Gln His
35 35 40 45
Pro Leu Ile Glu Ser Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln
50 55 60
Ser Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His
65 70 75 80
Tyr Tyr Trp Phe Leu Thr Asn Glu Gly Ile Asp Phe Leu Arg Thr Tyr
85 90 95
40 Leu Asn Leu Pro Ser Glu Ile Val Pro Ala Thr Leu Lys Lys Gln Gln
100 105 110
Lys Pro Leu Gly Arg Pro Phe Gly Gly Gly Asp Arg Pro Arg Gly
115 120 125
45 Pro Pro Arg Gly Asp Gly Glu Arg Arg Phe Gly Asp Arg Asp Gly Tyr
130 135 140
Arg Gly Gly Pro Lys Ser Gly Gly Glu Tyr Gly Asp Lys Ala Gly Ala
145 150 155 160
Pro Ala Asp Tyr Gln Pro Gly Phe Arg Gly Gly Ala Ser Gly Ala Arg
165 170 175
50 Gln Gly Phe Gly Arg Gly Ala Gly Gly Phe Gly Gly Gly Ala Gly Pro
180 185 190
Ala Ala Gly Ser Asp Leu Pro
195
(2) INFORMATION FOR SEQ ID NO:351:
55 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 183 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: peptide
60 (ix) FEATURE:
(A) NAME/KEY: peptide

```

496

(B) LOCATION: 1..183

(D) OTHER INFORMATION: / Ceres Seq. ID 1975986

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:351:

5 Met Ile Ile Ser Glu Thr Asn Arg Arg Glu Ile Ser Lys Tyr Leu Phe  
 1 5 10 15  
 Lys Glu Gly Val Leu Phe Ala Lys Lys Asp Phe Asn Leu Pro Gln His  
 20 25 30  
 Pro Leu Ile Glu Ser Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln  
 35 40 45  
 10 Ser Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His  
 50 55 60  
 Tyr Tyr Trp Phe Leu Thr Asn Glu Gly Ile Asp Phe Leu Arg Thr Tyr  
 65 70 75 80  
 Leu Asn Leu Pro Ser Glu Ile Val Pro Ala Thr Leu Lys Lys Gln Gln  
 85 90 95  
 15 Lys Pro Leu Gly Arg Pro Phe Gly Gly Gly Gly Asp Arg Pro Arg Gly  
 100 105 110  
 Pro Pro Arg Gly Asp Gly Glu Arg Arg Phe Gly Asp Arg Asp Gly Tyr  
 115 120 125  
 20 Arg Gly Gly Pro Lys Ser Gly Gly Glu Tyr Gly Asp Lys Ala Gly Ala  
 130 135 140  
 Pro Ala Asp Tyr Gln Pro Gly Phe Arg Gly Gly Ala Ser Gly Ala Arg  
 145 150 155 160  
 Gln Gly Phe Gly Arg Gly Ala Gly Gly Phe Gly Gly Gly Ala Gly Pro  
 165 170 175  
 25 Ala Ala Gly Ser Asp Leu Pro  
 180

(2) INFORMATION FOR SEQ ID NO:352:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1027 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..1027

(D) OTHER INFORMATION: / Ceres Seq. ID 1976019

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:352:

40 atcgaaatta gggtttcgcg ttaggagaag aagttaaagc aaaacacata caaacgcagt 60  
 caccttctct gtgccttcct tcttcaatct catcgcaatc atgatcatat ccgagactaa 120  
 tcgccgtgag atctccaagt acctcttcaa agaggggtgt ttgtttgcca aaaaggattt 180  
 caatttacca caacatcctt tgattgagag tgttccaaat ctgcaagtta tcaagttgat 240  
 gcagagtttc aaatctaagg aatatgtgag agagaccttt gcttgatgc attactactg 300  
 45 gttcctcaca aatgaaggta ttgactttct taggacttac cttaatctcc catctgagat 360  
 tgttcctgct actctgaaga agcaacagaa gcctcttggt cgaccttttg gaggtgggtg 420  
 tgaccgtccc cgtggccctc ctggtggtga tggagagagg aggtttggtg acagagatgg 480  
 ataccgtgga ggtcctaaat cagggtggaga gtatggtgac aaggctggag cacctgctga 540  
 ttaccagcct ggcttcaggg gtggagctag tggagcaagg caagggtttg gtcgtggagc 600  
 50 tgggtggtttt ggtggtggtg ctggtccagc tgctggatct gatctacctt gaaaaggaga 660  
 caatacagct cagccaaagc caaggccttt atcaccttac acaatgtacg cggacatgaa 720  
 gcctccaaca tcaccacttc catctccagt caccaatcat tagctatgtg agacgagaat 780  
 tggcctccta aaccctggag cttcttctta attgcaggtt gtaaacaatg agtaagagag 840  
 tgatggggca attcagtttt gcaggtatga atcagtgagc ttatttgtac aataatacaa 900  
 55 tattcatcaa catgccttat atatgagttc tacttttttt ctctgttca acgatcaatg 960  
 cagtaacat attgttacct tatctcatat atatatgatg acccaaatta attatcatta 1020  
 ggctttc

(2) INFORMATION FOR SEQ ID NO:353:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 183 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

497

(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..183

(D) OTHER INFORMATION: / Ceres Seq. ID 1976020

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:353:

```

Met Ile Ile Ser Glu Thr Asn Arg Arg Glu Ile Ser Lys Tyr Leu Phe
1      5      10      15
10  Lys Glu Gly Val Leu Phe Ala Lys Lys Asp Phe Asn Leu Pro Gln His
      20      25      30
    Pro Leu Ile Glu Ser Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln
      35      40      45
15  Ser Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His
      50      55      60
    Tyr Tyr Trp Phe Leu Thr Asn Glu Gly Ile Asp Phe Leu Arg Thr Tyr
      65      70      75      80
    Leu Asn Leu Pro Ser Glu Ile Val Pro Ala Thr Leu Lys Lys Gln Gln
      85      90      95
20  Lys Pro Leu Gly Arg Pro Phe Gly Gly Gly Gly Asp Arg Pro Arg Gly
      100      105      110
    Pro Pro Arg Gly Asp Gly Glu Arg Arg Phe Gly Asp Arg Asp Gly Tyr
      115      120      125
25  Arg Gly Gly Pro Lys Ser Gly Gly Glu Tyr Gly Asp Lys Ala Gly Ala
      130      135      140
    Pro Ala Asp Tyr Gln Pro Gly Phe Arg Gly Gly Ala Ser Gly Ala Arg
      145      150      155      160
    Gln Gly Phe Gly Arg Gly Ala Gly Gly Phe Gly Gly Gly Ala Gly Pro
      165      170      175
30  Ala Ala Gly Ser Asp Leu Pro
      180

```

(2) INFORMATION FOR SEQ ID NO:354:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 137 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..137

(D) OTHER INFORMATION: / Ceres Seq. ID 1976021

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:354:

```

Met Gln Ser Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp
1      5      10      15
45  Met His Tyr Tyr Trp Phe Leu Thr Asn Glu Gly Ile Asp Phe Leu Arg
      20      25      30
    Thr Tyr Leu Asn Leu Pro Ser Glu Ile Val Pro Ala Thr Leu Lys Lys
      35      40      45
50  Gln Gln Lys Pro Leu Gly Arg Pro Phe Gly Gly Gly Gly Asp Arg Pro
      50      55      60
    Arg Gly Pro Pro Arg Gly Asp Gly Glu Arg Arg Phe Gly Asp Arg Asp
      65      70      75      80
    Gly Tyr Arg Gly Gly Pro Lys Ser Gly Gly Glu Tyr Gly Asp Lys Ala
      85      90      95
55  Gly Ala Pro Ala Asp Tyr Gln Pro Gly Phe Arg Gly Gly Ala Ser Gly
      100      105      110
    Ala Arg Gln Gly Phe Gly Arg Gly Ala Gly Gly Phe Gly Gly Gly Ala
      115      120      125
60  Gly Pro Ala Ala Gly Ser Asp Leu Pro
      130      135

```

(2) INFORMATION FOR SEQ ID NO:355:

498

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 121 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..121

(D) OTHER INFORMATION: / Ceres Seq. ID 1976022

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:355:

```

Met His Tyr Tyr Trp Phe Leu Thr Asn Glu Gly Ile Asp Phe Leu Arg
1      5      10      15
Thr Tyr Leu Asn Leu Pro Ser Glu Ile Val Pro Ala Thr Leu Lys Lys
15     20     25     30
Gln Gln Lys Pro Leu Gly Arg Pro Phe Gly Gly Gly Gly Asp Arg Pro
35     40     45
Arg Gly Pro Pro Arg Gly Asp Gly Glu Arg Arg Phe Gly Asp Arg Asp
50     55     60
Gly Tyr Arg Gly Gly Pro Lys Ser Gly Gly Glu Tyr Gly Asp Lys Ala
20    65    70    75    80
Gly Ala Pro Ala Asp Tyr Gln Pro Gly Phe Arg Gly Gly Ala Ser Gly
85     90     95
Ala Arg Gln Gly Phe Gly Arg Gly Ala Gly Gly Phe Gly Gly Gly Ala
25   100   105   110
Gly Pro Ala Ala Gly Ser Asp Leu Pro
115   120

```

## (2) INFORMATION FOR SEQ ID NO:356:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 478 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..478

(D) OTHER INFORMATION: / Ceres Seq. ID 1976673

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:356:

```

aaattcaatc tcttccaatt tctcttcttc ttctctcgca cgcgttaatg gcggttcgt      60
tctcactcac gagcttcac tccttcatct caccattcaa atctcaaacc aaacctacac      120
caccacaaaa tctcactctt ccttctccaa ctatctccca aaggcgaaga aatgatctcg      180
ctatcgaatc aatggcggtc gaagaatctt cttcaaccgc ttcttcaactt tcctctgagc      240
ttgcttctgt gatatgcccc tcgcttgctt actccaacac gctcttcttc agttctggat      300
acaatgtgca agtgtttgtt gaagataacg agtcagagga gaggttggtg aatcgattta      360
ggagagaagt gatgagaact ggtgttatac aggaatgtaa gaggagaaga tactttgaga      420
ataaacaaga tgagaagaaa cgtaggactc gtgatgctgc taagcgtaat aagaaaag

```

## (2) INFORMATION FOR SEQ ID NO:357:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 158 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..158

(D) OTHER INFORMATION: / Ceres Seq. ID 1976674

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:357:

```

Ile Gln Ser Leu Pro Ile Ser Leu Leu Pro Arg Asp Ala Leu Met
1      5      10      15
Ala Ala Ser Phe Ser Leu Thr Ser Phe Ile Ser Phe Ile Ser Pro Phe

```

5 Lys Ser Gln Thr Lys Pro Thr Pro Pro Asn Leu Thr Leu Pro Ser 30  
 35 Pro Thr Ile Ser Gln Arg Arg Asn Asp Leu Ala Ile Glu Ser Met 45  
 50 Ala Val Glu Glu Ser Ser Thr Ala Ser Ser Ser Ser Glu Leu 60  
 65 Ala Ser Val Ile Cys Pro Ser Leu Ala Tyr Ser Asn Thr Leu Phe 75  
 80 Ala Ser Val Ile Cys Pro Ser Leu Ala Tyr Ser Asn Thr Leu Phe 85  
 90 Ser Ser Gly Tyr Asn Val Gln Val Phe Val Glu Asp Asn Glu Ser Glu 95  
 100 Glu Arg Leu Val Asn Arg Phe Arg Arg Glu Val Met Arg Thr Gly Val 110  
 115 Ile Gln Glu Cys Lys Arg Arg Arg Tyr Phe Glu Asn Lys Gln Asp Glu 125  
 130 Lys Lys Arg Arg Thr Arg Asp Ala Ala Lys Arg Asn Lys Lys 140  
 145 (1) INFORMATION FOR SEQ ID NO:359:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 143 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: linear  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..143  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1976675  
 20 Met Ala Ala Ser Phe Ser Leu Thr Ser Phe Ile Ser Phe Ile Ser Pro 15  
 1 Phe Lys Ser Gln Thr Lys Pro Thr Pro Pro Asn Leu Thr Leu Pro 30  
 25 Ser Pro Thr Ile Ser Gln Arg Arg Asn Asp Leu Ala Ile Glu Ser 45  
 35 Met Ala Val Glu Glu Ser Ser Thr Ala Ser Ser Ser Ser Glu 60  
 50 Leu Ala Ser Val Ile Cys Pro Ser Ser Leu Ala Tyr Ser Asn Thr Leu Phe 75  
 80 Phe Ser Ser Gly Tyr Asn Val Gln Val Phe Val Glu Asp Asn Glu Ser 95  
 90 Glu Glu Arg Leu Val Asn Arg Phe Arg Arg Glu Val Met Arg Thr Gly 105  
 110 Val Ile Gln Glu Cys Lys Arg Arg Arg Tyr Phe Glu Asn Lys Gln Asp 125  
 130 Glu Lys Lys Arg Arg Thr Arg Asp Ala Ala Lys Arg Asn Lys Lys 140  
 145 (2) INFORMATION FOR SEQ ID NO:359:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 95 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: linear  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..95  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1976676  
 55 Met Ala Val Glu Ser Ser Ser Thr Ala Ser Ser Ser Ser Ser Glu 15  
 1 Leu Ala Ser Val Ile Cys Pro Ser Ser Leu Ala Tyr Ser Asn Thr Leu Phe 30

20 Phe Ser Ser Gly Tyr Asn Val Gln Val Phe Val Glu Asp Asn Glu Ser 30  
 35 Glu Glu Arg Leu Val Asn Arg Phe Arg Arg Glu Val Met Arg Thr Gly 45  
 50 Val Ile Gln Glu Cys Lys Arg Arg Tyr Phe Glu Asn Lys Gln Asp 60  
 65 Glu Lys Lys Arg Arg Thr Arg Asp Ala Ala Lys Arg Asn Lys Lys 75  
 80 (2) INFORMATION FOR SEQ ID NO:360:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1076 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..1076  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2025186  
 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:360:  
 aatcattgc agagaaaaa taataaaaaa aaacagagaa aagagatat ggaatcaatc 60  
 ggagtcctta tcatgtccc catgtctcc taccctgaga acgacttga gaagcgttc 120  
 aaactcttc gttctggac tctctcgag aaactcgtcc tctagaaac tcatcgaaac 180  
 tccatccgcg ccgtgtttgg gaatgcttt gccggcgccg atgtcagat catcagcgt 240  
 ctgccaaac ttagattgt attcagttc agctcggcg tcacaagat cgaattgggg 300  
 aatgcaag aaaaggat ccgctcacc aaacccccc acgtctcac cgaatggc 360  
 gccttagt gccatcaa ttacttca agaaccatta agctcagt cgcctcacag 420  
 tatctcca cggtgttga cctgtcca aaactcaga tctctgtt cgtatgccg 480  
 ttgacgagc agaccagaca cattgtgac cgcagatca tgaatcatt agagctaa 540  
 ggcctctca taacattgg ccgtggaca catgtgat agcaagct tattaaagt 600  
 ctacagaa ccgcttagg tgggctgcc cttagtgt ttgacagga gccacagt 660  
 ccgagagc tcttgacct tgaagtta gtctcttc ctacgttg gactggcact 720  
 gggaaacac ggaatgccat ggcgatctt gtcggtgta acttgaagc gcaattct 780  
 cttcattga aagtttita tgttaacga ctcataaaa gggtaattt tcttcc 840  
 900 1020  
 (2) INFORMATION FOR SEQ ID NO:361:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 329 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: linear  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..329  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2025187  
 45 An His Leu Gln Lys Lys Asn Lys Lys Lys Lys Gln Arg Lys Gly Asp 15  
 1 Met Glu Ser Ile Gly Val Leu Met Met Cys Pro Met Ser Ser Tyr Leu 30  
 20 Glu Asn Glu Leu Glu Lys Arg Phe Asn Leu Leu Arg Phe Thr Ser 35  
 40 Pro Glu Lys Ser Val Leu Leu Glu Thr His Arg Asn Ser Ile Arg Ala 45  
 50 Val Val Gly Asn Ala Ser Ala Gly Ala Asp Ala Gln Leu Ile Ser Asp 55  
 60 Leu Pro Asn Leu Glu Ile Val Ser Ser Phe Ser Val Gly Leu Asp Lys 80

```

502
150
155
160
Lys Arg Ala Glu Ala Phe Ser Cys Pro Ile Asn Tyr Tyr Ser Arg Thr
165
170
175
Ile Lys Pro Asp Val Ala Tyr Lys Tyr Pro Thr Val Val Asp Leu
180
185
190
Ala Gln Asn Ser Asp Ile Leu Val Val Ala Cys Pro Leu Thr Glu Gln
195
200
205
Thr Arg His Ile Val Asp Arg Gln Val Met Asp Ala Leu Gly Ala Lys
210
215
220
Gly Val Leu Ile Asn Ile Gly Arg Gly Pro His Val Asp Glu Gln Glu
225
230
235
Leu Ile Lys Ala Leu Thr Glu Gly Arg Leu Gly Gly Ala Ala Leu Asp
240
245
250
255
Val Phe Glu Gln Glu Pro His Val Pro Glu Glu Leu Phe Gly Leu Glu
260
265
270
Asn Val Val Leu Leu Pro His Val Gly Ser Gly Thr Val Glu Thr Arg
275
280
285
Asn Ala Met Ala Asp Leu Val Val Gly Asn Leu Glu Ala His Phe Ser
290
295
300
Gly Lys Ser Leu Leu Thr Pro Val Val
310
315
320
325
330
335
340
345
350
355
360
365
370
375
380
385
390
395
400
405
410
415
420
425
430
435
440
445
450
455
460
465
470
475
480
485
490
495
500
505
510
515
520
525
530
535
540
545
550
555
560
565
570
575
580
585
590
595
600
605
610
615
620
625
630
635
640
645
650
655
660
665
670
675
680
685
690
695
700
705
710
715
720
725
730
735
740
745
750
755
760
765
770
775
780
785
790
795
800
805
810
815
820
825
830
835
840
845
850
855
860
865
870
875
880
885
890
895
900
905
910
915
920
925
930
935
940
945
950
955
960
965
970
975
980
985
990
995

```

[illegible]



- 225 503 230 235 240  
Gly Arg Leu Gly Gly Ala Ala Leu Asp Val Phe Glu Gln Glu Pro His 255  
245 250  
Val Pro Glu Glu Leu Phe Gly Leu Glu Asn Val Val Leu Leu Pro His 270  
260 265 270  
Val Gly Ser Gly Thr Val Glu Thr Arg Asn Ala Met Ala Asp Leu Val 285  
275 280 285  
Val Gly Asn Leu Glu Ala His Phe Ser Gly Lys Ser Leu Leu Thr Pro 300  
290 295 300  
Val Val 305
- (2) INFORMATION FOR SEQ ID NO:364:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 555 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..555  
(D) OTHER INFORMATION: / Ceres Seq. ID 2025372  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:364:  
aaactcact ttcacttga caagataag gaacacatg ctgtgtcagc gatattggt 60  
accggaatcg tcaactgtgc tgetctcgc gttctccgcg aatttcaagt tccaaaattg 120  
ggtaattggg gaggatcagg gatgttgata gaggttcgt cgaggccaca gaagaaatcg 180  
acagatcacc acaggaagac gaggcggaag aagactcagc ctggggaacat taagagaag 240  
ctacatggtt accgtctcct tctctctctt ccggcggaat ggagtcagtt cactctcgt 300  
ctgacgacg ggtgtgtcgc cactgtcgc ggagattgg ttccaggcgc tgcctagttg 360  
gtargagta tctgcgatt tggttgtaat ctagtttag aactttgtt gttgttact 420  
ggttttcag ttcttgatg ttgtgtgtg ggtttgttg tgttgatg ggaatgaat 480  
ggaccaaat tgcgattat aagacttcaa cttctctct cgtttttga gtttaaggt 540  
ctcaacttta tagtt
- (2) INFORMATION FOR SEQ ID NO:365:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 118 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..118  
(D) OTHER INFORMATION: / Ceres Seq. ID 2025373  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:365:  
Lys Thr His Phe His Leu His Lys Asp Lys Glu Thr Met Ser Val Ser 15  
1 5 10 15  
Ala Ile Phe Gly Thr Gly Ile Val Thr Val Ala Ala Ser Pro Val Leu 20  
20 25 30 35  
Arg Gln Phe Gln Val Pro Lys Leu Gly Asn Gly Gly Gly Leu Gly Met 40  
35 40 45  
Val Ile Glu Cys Ser Ser Arg Pro Gln Lys Lys Ser Thr Ala His 50  
50 55 60  
Arg Lys Thr Arg Pro Lys Lys Thr Gln Pro Trp Asp Ile Lys Arg Lys 80  
65 70 75 80  
Pro Thr Val Tyr Ala Pro Leu Pro Pro Leu Pro Ala Glu Trp Ser Pro 95  
85 90  
Phe Thr Leu Ala Ser Asp Gly Gly Ala Ala Thr Ala Ala Gly Asp 100  
105 110  
Leu Val Ser Gly Ala Ala 115
- (2) INFORMATION FOR SEQ ID NO:366:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 106 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..106  
(D) OTHER INFORMATION: / Ceres Seq. ID 2025374  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:366:  
Met Ser Val Ser Ala Ile Phe Gly Thr Gly Ile Val Thr Val Ala Ala 15  
1 5 10 15  
Ser Pro Val Leu Arg Gln Phe Gln Val Pro Lys Leu Gly Asn Gly Gly 30  
35 40 45  
Gly Leu Gly Met Val Ile Glu Cys Ser Ser Arg Pro Gln Lys Lys Ser 50  
55 60  
Thr Ala His His Arg Lys Thr Arg Pro Lys Lys Thr Gln Pro Trp Asp 65  
70 75 80  
Ile Lys Arg Lys Pro Thr Val Tyr Ala Pro Leu Pro Pro Leu Pro Ala 85  
80 85  
Glu Trp Ser Pro Phe Thr Leu Ala Ser Asp Asp Gly Gly Ala Ala Thr 95  
90 95  
Ala Ala Gly Asp Leu Val Ser Gly Ala Ala 100  
105 110  
(2) INFORMATION FOR SEQ ID NO:367:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 71 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..71  
(D) OTHER INFORMATION: / Ceres Seq. ID 2025375  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:367:  
Met Val Ile Glu Cys Ser Ser Arg Pro Gln Lys Lys Ser Thr Ala His 15  
1 5 10 15  
His Arg Lys Thr Arg Pro Lys Lys Thr Gln Pro Trp Asp Ile Lys Arg 20  
20 25 30 35  
Lys Pro Thr Val Tyr Ala Pro Leu Pro Pro Leu Pro Ala Glu Trp Ser 40  
35 40 45  
Pro Phe Thr Leu Ala Ser Asp Asp Gly Gly Ala Ala Thr Ala Ala Gly 60  
50 55 60  
Asp Leu Val Ser Gly Ala Ala 70  
65 70  
(2) INFORMATION FOR SEQ ID NO:368:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 631 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..631  
(D) OTHER INFORMATION: / Ceres Seq. ID 2025471  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:368:  
atgtcgcaaa agttcttaaa accgcaact gtcagtcact tatccaacag taatccaatt 60  
attgtcaag cctcgggtt attaccgtt aaaccgaaa gaaaaaactt cttcggaaat 120  
ctgaagaacc ataatcagc agaaaaaat cttaaggatc cttaagttta taatcggatg 180

505

ataaagatga ttaagtcagtt ttccggttaag ttgatgaagg agaagagttg taccgggtga  
 240  
 agccggagga aataagactt aagatctgcc gtcaagaagt cggctctcgtc tccacagagt  
 300  
 ggcctcgtc cgtcgtcgtt gagagagctt aaggagatc tgaatcagag tcatgttgtt  
 360  
 gggcggtga gtagagagtt gagcaagcg gagaaattc ttgagagagtt gatgtttcgt  
 420  
 agctgttggg gatcttcta gatctgttgc agaaataat agatagagaa acgaaccaa  
 480  
 acggttcag ttatgtttt ttgattttgg atttttgtt ttctcgtgaa tagttttttt  
 540  
 tttttttc gattttttg ttatttgtgt ttatgtana tcatgtataa aatgaaagt  
 600  
 aaatacga agataagtt ttgtcttcac c

(2) INFORMATION FOR SEQ ID NO:369:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 146 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ix) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..146

(D) OTHER INFORMATION: / Ceres Seq. ID 2025472

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:369:

Met Ser Arg Lys Phe Leu Lys Pro Pro Arg Val Ser His Leu Ser Asn  
 1 5 10 15  
 Ser Asn Pro Ile Ile Arg Gln Ala Ser Gly Leu Leu Pro Ser Lys Pro  
 20 25 30

Lys Arg Lys Asn Ser Leu Arg Asn Leu Lys Ser His Lys Ser Ala Glu  
 35 40 45  
 Lys Asn Leu Lys Asp Arg Lys Phe Tyr Asn Arg Met Asn Arg Val Ile  
 50 55 60

Ser Gln Phe Ser Gly Lys Leu Met Lys Glu Lys Ser Val Thr Gly Val  
 65 70 75 80  
 Ser Gly Gly Asp Thr Thr Leu Arg Ser Ala Val Lys Glu Ser Val Ser  
 85 90 95

Ser Pro Gln Ser Ala Ser Ser Ser Ser Val Arg Arg Leu Lys Gly  
 100 105 110  
 Asp Leu Glu Ser Ser Arg Phe Gly Ala Ala Ser Glu Arg Leu Arg  
 115 120 125

Gln Ala Glu Glu Ser Leu Arg Thr Val Met Phe Leu Ser Cys Trp Gly  
 130 135 140  
 Ser Cys

(2) INFORMATION FOR SEQ ID NO:370:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 87 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ix) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..87

(D) OTHER INFORMATION: / Ceres Seq. ID 2025473

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:370:

Met Asn Arg Val Ile Ser Gln Phe Ser Gly Lys Leu Met Lys Glu Lys  
 1 5 10 15  
 Ser Val Thr Gly Val Ser Gly Gly Asp Thr Thr Leu Arg Ser Ala Val  
 20 25 30

Lys Glu Ser Val Ser Pro Gln Ser Ala Ser Ser Ser Ser Val  
 35 40 45  
 Arg Arg Leu Lys Gly Asp Leu Glu Ser Ser Arg Phe Gly Ala Ala Ala  
 50 55 60

Ser Glu Arg Leu Arg Gln Ala Glu Ser Leu Arg Thr Val Met Phe  
 65 70 75 80

506

Leu Ser Cys Trp Gly Ser Cys  
 85

(2) INFORMATION FOR SEQ ID NO:371:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 75 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ix) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..75

(D) OTHER INFORMATION: / Ceres Seq. ID 2025474

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:371:

Met Lys Glu Lys Ser Val Thr Gly Val Ser Gly Gly Asp Thr Thr Leu  
 1 5 10 15  
 Arg Ser Ala Val Lys Glu Ser Val Ser Ser Pro Gln Ser Ala Ser Ser  
 20 25 30

Ser Ser Ser Val Arg Arg Leu Lys Gly Asp Leu Glu Ser Ser Arg Phe  
 35 40 45  
 Gly Ala Ala Ala Ser Glu Arg Leu Arg Gln Ala Glu Ser Leu Arg  
 50 55 60

Thr Val Met Phe Leu Ser Cys Trp Gly Ser Cys  
 65 70 75

(2) INFORMATION FOR SEQ ID NO:372:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2029 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..2029

(D) OTHER INFORMATION: / Ceres Seq. ID 2025475

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:372:

atgggtgaaga gtctgtttc agctgtgaag aaagcattaa gccacgaacc aaacacaaa  
 60  
 aaagacaga agccacataa gtccaaagaa tgggttgga atccaaagaa gctagtgtt  
 120  
 actaatctg gtgcagcata ttctctcgt actgtcaag acgcgaact aaagagatt  
 180  
 gagagcaac agacagaca tgcctactct gtgctattg caactgtgc agtgcagag  
 240  
 gcagcgttg cagctgcta agtgcgtct gaagtgttc gctctctgc attatcacg  
 300  
 ttcccggga aatcaatga agatgcgcc gctatcaaga ttacagacga attacagga  
 360  
 tatatggaa gaagcatt gcgtgcttg agagctttg tgaagctaaa atctttagt  
 420  
 cagggaat gtgtgagcg tcaagcaca tctacattgc aaagctaga aacattagt  
 480  
 agagcacat atcagattcg tgaggaaggg cctcattgt ctgaggataa acaggttta  
 540  
 acacgacgc tccacaaa acacataaa gactttgata agactggga aatttgaat  
 600  
 gtagtacct tgtcgggga gaaagtga gactgtgca agcgtgcca tgcctgagat gattccaga  
 660  
 atgagaagag aaaaagcgt tgcattgca ttacgtcacc agaatcacg gaaaactca  
 720  
 actaaaatg gttcctaac attcattgac ctaaacatc cgtattggg ttgagtgtg  
 780  
 ctagagactt gtaggtgc tgcaccaac gaaaacact cactcaccc agataatgt  
 840  
 gaaaagact cgttgcgtg gactgtgca agcgtgcca tgcctgagat gattccaga  
 900  
 ggcacaaacc ttccacaa aggaagaca ccaaacgtc gaaggggtc agcccgaga  
 960  
 tgaggaag tcccaatga agactcaaac agcattgtga gtttcaact agaacact  
 1020  
 tgaatcgtt ggcatac tttgtgata attccatcaa ctacacaga cgcacagca  
 1080  
 accatgatt tctctcagt agttccagg taccatgac ctacacaga cgcacagca  
 1140  
 agctgcatt tctcaacct tagtctcta agctcagca agacagcga aaacagctt  
 1200  
 tctcttcgg agtcacac gactgaaga cggttttcg tagggattct agggcttcc  
 1260  
 ttcatctgg agtcacac gactttaca tctcttctt ctctctctc gatattctc  
 1320  
 agtgtgta ttctgggaa aactttgga gaaaagagc agaaatgag cggaccgga  
 1380  
 agaaatcgc gga-gtggc ttcaagctt cgaagactat ccatgggat agtatggta  
 1440  
 aggtcttgt cacagttag gctctagag agttctcaa ccttcgctg gctttcagt  
 1500  
 aggttaaac acagctcag accaaatta gtcaggaacc tgaaccata gattggatt  
 1560



509	Ala Lys Ala Arg Ala Arg Phe Ser Asn Leu Ser Pro Leu Ser Ser Glu	275	Val Cys Ile Leu Gly Lys Thr Leu	325	510
275	Lys Thr Ala Lys Lys Arg Leu Ser Phe Ser Gly Pro Lys Thr Val	285			
290	Arg Arg Phe Ser Val Gly Ile Leu Gly Leu Ser Phe Ile Gly Gln Ile	295			
305	Arg Arg Phe Thr Ser Ser Ser Ser Ser Ile Phe Val Ser	310			
325	Val Cys Ile Leu Gly Lys Thr Leu	330			
340					
10	(2) INFORMATION FOR SEQ ID NO:375:				
	(i) SEQUENCE CHARACTERISTICS:				
	(A) LENGTH: 328 amino acids				
	(B) TYPE: amino acid				
	(C) STRANDEDNESS:				
	(D) TOPOLOGY: linear				
	(ii) MOLECULE TYPE: peptide				
	(ix) FEATURE:				
	(A) NAME/KEY: peptide				
	(B) LOCATION: 1..328				
	(D) OTHER INFORMATION: / Ceres Seq. ID 2025478				
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:375:				
	Met Ala Arg Arg Ala Leu Arg Ala Leu Arg Gly Leu Val Arg Leu Lys	10			
	1 5 10 15				
25	Ser Leu Val Gln Gly Lys Cys Val Arg Arg Gln Ala Thr Ser Thr Leu	20			
	Gln Ser Met Gln Thr Leu Ala Arg Val Gln Tyr Gln Ile Arg Glu Arg	25			
	35 40 45				
30	Arg Leu Arg Leu Ser Gln Asp Lys Gln Ala Leu Thr Arg Gln Leu Gln	50			
	Gln Lys His Asn Lys Asp Phe Asp Lys Thr Gly Glu Asn Trp Asn Asp	55			
	65 70 75 80				
	Ser Thr Leu Ser Arg Glu Lys Val Glu Ala Asn Met Leu Asn Lys Gln	85			
	90 95				
35	Val Ala Thr Met Arg Arg Glu Lys Ala Leu Ala Tyr Ala Phe Ser His	100			
	Gln Asn Thr Trp Lys Asn Ser Thr Lys Met Gly Ser Gln Thr Phe Met	105			
	115 120 125				
40	Asp Pro Asn Asn Pro His Trp Gly Trp Ser Trp Leu Glu Arg Trp Met	130			
	Ala Ala Arg Pro Asn Glu Asn His Ser Leu Thr Pro Asp Asn Ala Glu	135			
	145 150 155				
	Lys Asp Ser Ser Ala Arg Ser Val Ala Ser Arg Ala Met Ser Glu Met	165			
	170 175				
45	Ile Pro Arg Gly Lys Asn Leu Ser Pro Arg Gly Lys Thr Pro Asn Ser	180			
	Arg Arg Gly Ser Ser Pro Arg Val Arg Gln Val Pro Ser Glu Asp Ser	185			
	195 200 205				
50	Asn Ser Ile Val Ser Phe Gln Ser Glu Gln Pro Cys Asn Arg Arg His	210			
	Ser Thr Cys Gly Ser Ile Pro Ser Thr Arg Asp Asp Glu Ser Phe Thr	215			
	225 230 235				
	Ser Ser Phe Ser Gln Ser Val Pro Gly Tyr Met Ala Pro Thr Gln Ala	240			
	245 250 255				
55	Ala Lys Ala Arg Ala Arg Phe Ser Asn Leu Ser Pro Leu Ser Ser Glu	260			
	Lys Thr Ala Lys Lys Arg Leu Ser Phe Ser Gly Ser Pro Lys Thr Val	265			
	275 280 285				
	Arg Arg Phe Ser Val Gly Ile Leu Gly Leu Ser Phe Ile Gly Gln Ile	290			
	295 300				
60	Arg Arg Arg Phe Thr Ser Ser Ser Ser Ser Ile Phe Val Ser	305			
	310 315				
	(2) INFORMATION FOR SEQ ID NO:376:				
	(i) SEQUENCE CHARACTERISTICS:				
	(A) LENGTH: 678 base pairs				
	(B) TYPE: nucleic acid				
	(C) STRANDEDNESS: single				
	(D) TOPOLOGY: linear				
	(ii) MOLECULE TYPE: DNA (genomic)				
	(ix) FEATURE:				
	(A) NAME/KEY: -				
	(B) LOCATION: 1..678				
	(D) OTHER INFORMATION: / Ceres Seq. ID 2025524				
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:376:				
	gatttttaag ggtttcaagt gaaagagata atagcgccg gaggccatgg ttctcaagac	30			
	tgagctttgc cgatttcagt gcagaaaaat ttacctggt agaggatca gattttccg	35			
	atcggaactt cagggttttt tgtttctcaa ctccaaatgt aagagttatt tccacaacaa	40			
	gttgaagcca tctaaagtttt gctggactgc tatgtaccga aagcagcaca aagaggccgc	45			
	agcacaaag gctgtgaaga gaaggaagcg tgcactaag aagctttact caagttcgat	50			
	tgtcgtgct actttggagg ttattcagaa gaagcagaca gagaagctg aagttctga	55			
	tgcgctaga gaagctgcc tactgtagat caagcagaca atcaagaaga ccaaggacga	60			
	gaagaagga aagaaggtcgc agtatgcata aagaacacag aagtcacaa gaaaggaaaa	65			
	tatccccaa agtgtgtcac ccaaggtcgc taagtgggt ggtgtggag ggaagcttg	70			
	aatggagta tagatgacc cactttctc ttctcacta tctttcttc ttgtttgac	75			
	atgttttgt ttgtcagcc atttttaagt ttgtcaccag atctaataa ttcaatttat	80			
	gaaaccttt tgtttggc	85			
	(2) INFORMATION FOR SEQ ID NO:377:				
	(i) SEQUENCE CHARACTERISTICS:				
	(A) LENGTH: 164 amino acids				
	(B) TYPE: amino acid				
	(C) STRANDEDNESS:				
	(D) TOPOLOGY: linear				
	(ii) MOLECULE TYPE: peptide				
	(ix) FEATURE:				
	(A) NAME/KEY: peptide				
	(B) LOCATION: 1..164				
	(D) OTHER INFORMATION: / Ceres Seq. ID 2025525				
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:377:				
	Met Val Leu Lys Thr Glu Leu Cys Arg Phe Ser Gly Gln Lys Ile Tyr	1			
	1 5 10 15				
40	Pro Gly Arg Gly Ile Arg Phe Ile Arg Ser Asp Ser Gln Val Phe Leu	20			
	Phe Leu Asn Ser Lys Cys Lys Arg Tyr Phe His Asn Lys Leu Lys Pro	25			
	35 40 45				
45	Ser Lys Leu Cys Trp Thr Ala Met Tyr Arg Lys Gln His Lys Lys Asp	50			
	Ala Ala Gln Glu Ala Val Lys Arg Arg Arg Ala Thr Lys Lys Pro	55			
	65 70 75 80				
	Tyr Ser Arg Ser Ile Val Gly Ala Thr Leu Glu Val Ile Gln Lys Lys	85			
	Arg Ala Glu Lys Pro Glu Val Arg Asp Ala Ala Arg Glu Ala Ala Leu	90			
	100 105 110				
50	Arg Glu Ile Lys Glu Arg Ile Lys Lys Thr Lys Asp Glu Lys Lys Ala	115			
	Lys Lys Val Glu Tyr Ala Ser Lys Gln Gln Lys Ser Gln Val Lys Gly	120			
	130 135 140				
	Asn Ile Pro Lys Ser Ala Ala Pro Lys Ala Ala Lys Met Gly Gly	145			
	Gly Gly Arg Arg	150 155			
60	(2) INFORMATION FOR SEQ ID NO:378:				
	(i) SEQUENCE CHARACTERISTICS:				

511

(A) LENGTH: 109 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..109  
(D) OTHER INFORMATION: / Ceres Seq. ID 2025526  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:378:  
Met Tyr Arg Lys Gln His Lys Lys Asp Ala Ala Gln Glu Ala Val Lys  
1 5 10 15  
Arg Arg Arg Arg Ala Thr Lys Lys Pro Tyr Ser Arg Ser Ile Val Gly  
20 25 30 35 40 45  
Ala Thr Leu Glu Val Ile Gln Lys Lys Arg Ala Glu Lys Pro Glu Val  
50 55 60  
Arg Asp Ala Ala Arg Glu Ala Ala Leu Arg Glu Ile Lys Glu Arg Ile  
65 70 75 80  
Lys Lys Thr Lys Asp Glu Lys Lys Ala Lys Lys Val Glu Tyr Ala Ser  
85 90 95  
Lys Gln Gln Lys Ser Gln Val Lys Gly Asn Ile Pro Lys Ser Ala Ala  
100 105  
Pro Lys Ala Ala Lys Met Gly Gly Gly Arg Arg  
(2) INFORMATION FOR SEQ ID NO:379:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1623 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..1623  
(D) OTHER INFORMATION: / Ceres Seq. ID 2025544  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:379:

atggcgatt ctaagatgt tctctgaaa tcaacatgt ttgtttttt ctgattact  
120 tgcgagat tcaaatgtc aagaacgta aacagactt tctgggttc attcacttc  
180 attctctt cgcgtatgt tgggtagcc gccattgtg ggcacatcga aggtcaccc  
240 actgacatc gattcttca agtagtgtt ctggagtgg tgcgtgagt tatcactgg  
300 gctcaactt tggaccggt gclacatgt gatcaacct tctcaaggt gcttacttc  
360 aggaagtag tgaatggaa agccattat ggaattggtta gacttttgt tctcaagca  
420 tatcaatgc aataatagc attgataca agttacatg agagtcaaa attatcgat  
480 tcaatcatg aaaaaaagg acatcgaag agctctatc agaacatccc gatgtctac  
540 aaccgttcag aacatcaac aacatcgaat tgcgtgatt gcttacagga tgggaagaa  
600 gggaagtag gaagaagct agcaatggt ggcacacat tctatgaa ttgcatagt  
660 gactgttc tgaacaaga aactgtccc attgaagaa gacaagaagaa gaagcacag  
720 aatctctc ttaaaatcca atgcatacga ctacgtctc tctctctc tcaactact  
780 tctctctc cttctcgtt cttctcgtc tctatctca tctatctct ctgctgcgg  
840 tgaatgact tgcgtgatt tcaatcctat aaatactca aactcagcc tctcaccgc  
900 actctccc actctaccc cactgtgctc tgcgtgtgc ggtgcgcaa catctcgtg  
960 tctctctc tcaagaatgt cttctagat tccctctc tccctctc ctgtctctc  
1020 tctctctg aagaaggaag tctgtctc tgaagaagca ccaagtctt tgggaacta  
1080 cttcagccc ataaagca ataatgtt tttcttca ggtgtttg gacttacc  
1140 tgaactgga aagtgttct cgaagatgt cgaagatcag actgagcag tactcaaaa  
1200 catggggag atattgaag ctgaggtgc tgaattcc tgggtgtga agacaacat  
1260 catgttggc gatttggct actcaagc agtgaacag atatatcca aactctccc  
1320 agctctctt cagcagcat cagctatca agttgact tgccttca acgcgaagt  
1380 cgaagttaa tgaatgcaa cactcaga cactcaaat aactctac aggggtatt  
1440 tggccgaa gatgtcaaa caaatgaag aaacaatgt ttttgggtc taaccatat  
1500 ttaactact tttactgt tttactacg aaaaattgt cggattacc accattaggt  
1560 tggttacaa tcaaaagctt ctactcaaga aaagtatt tagttctt tccacctca

512

tttgtgttg acgtatatgt attgattgta tacattattt tatttgagtg aaaccattat  
gct  
(2) INFORMATION FOR SEQ ID NO:380:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 249 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..249  
(D) OTHER INFORMATION: / Ceres Seq. ID 2025545  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:380:  
Met Ser Gly Cys Leu Asp Lys Lys Leu Ala Pro Phe Glu Glu Asp Lys  
1 5 10 15  
Arg Arg Ser Thr Arg Ile Ile Cys Leu Lys Ser Asn Arg Ile Asp Ser  
20 25 30 35 40 45  
Val Phe Phe Phe Phe Thr Tyr Thr Ser Leu Ser Leu Ser Trp Ser  
50 55 60  
Ile Arg Tyr Ser Ile Leu Leu Val Gly Ala Gly Glu Ser Glu Met Thr  
65 70 75 80  
Trp Ser Val Phe Arg Ser Ile Asn Thr Pro Thr Leu Asp Leu Ser Thr  
85 90 95  
Ala Thr Phe Ala Gly Val Ser Leu Phe Arg Met Ser Ser Arg Ser Pro  
100 105 110  
Pro Phe Ala Ser Leu Ser Val Ser Ala Ser Ser Val Lys Lys Glu Val  
115 120 125  
Val Ser Thr Glu Lys Ala Pro Ala Ala Leu Gly Pro Tyr Ser Gln Ala  
130 135 140  
Ile Lys Ala Asn Asn Leu Val Phe Leu Ser Gly Val Leu Gly Leu Ile  
145 150 155 160  
Pro Glu Thr Gly Lys Phe Val Ser Glu Ser Val Glu Asp Gln Thr Glu  
165 170 175  
Gln Val Leu Lys Asn Met Gly Glu Ile Leu Lys Ala Ser Gly Ala Asp  
180 185 190  
Tyr Ser Ser Val Val Lys Thr Thr Thr Ile Met Leu Ala Asp Ala Asp  
195 200 205  
Phe Lys Thr Val Asn Glu Ile Tyr Ala Lys Tyr Phe Pro Ala Pro Ser  
210 215 220  
Pro Ala Arg Ser Thr Tyr Gln Val Ala Ala Leu Pro Leu Asn Ala Lys  
225 230 235 240  
Ile Glu Ile Glu Cys Ile Ala Thr Leu  
245  
(2) INFORMATION FOR SEQ ID NO:381:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 211 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..211  
(D) OTHER INFORMATION: / Ceres Seq. ID 2025546  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:381:  
Met Gly Ile Ser Lys Met Ile Leu Leu Lys Phe Thr Tyr Val Val Ser  
1 5 10 15  
Phe Trp Ile Thr Cys Val Ile Phe Lys Met Ser Arg Thr Ile Lys Arg  
20 25 30

5	Ala Leu Leu Gly Ser Phe Ile Leu Ala Ser Ala Ser Val Val	513	(2) INFORMATION FOR SEQ ID NO:383:
	35 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205	514	(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 542 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: DNA (genomic) (ix) FEATURE: (A) NAME/KEY: - (B) LOCATION: 1..542 (D) OTHER INFORMATION: / Ceres Seq. ID 2026207 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:383: tcacacaaat cctctctc tcttcattg tctctctta atggaaacaa ccgagaatc targtacta gaactacta gtaccacaa gaggagacc aaatcgaag atagtctcc agttatgcy gagaagatg atgtgaaga gttgtatca gattatgca aggtttcag ttgtttgcy gatccagga gactctcat cacagtctg tcttaagc gaactcagg gatacttga attgaagga tgagaagga agatctcaa ggaatggtta gagaagaga ccctatgga gatgtgttc ttacacaa cgaattctg gtctctagg tctgttaag ccctgatg atggagagc ccgaacttg gttgagaa gcaactacc aagaacttg taatcacaat cctctctta tgccttgat gatctctc cttcttggt tatctctt ttactctac aaacctata aagtctctg aatataaat gaagtaatga ttcttgatc cc
10	Val Ala Leu Leu Arg Val Val Asn Gly Lys Ala Ile Met Gly Leu	10	(2) INFORMATION FOR SEQ ID NO:384:
15	Val Arg Pro Phe Val Leu Lys Ala Tyr Gln Trp Gln Ile Ile Ala Leu	15	(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 148 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide (ix) FEATURE: (A) NAME/KEY: poptide (B) LOCATION: 1..148 (D) OTHER INFORMATION: / Ceres Seq. ID 2026208 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:384: His Gln Asn Leu Phe Ser Leu Phe His Cys Leu Pro Leu Met Glu Pro 1 5 10 15 Thr Glu Lys Ser Met Leu Leu Glu Thr Thr Ser Thr Thr Lys Met Glu 20 25 30 Thr Lys Tyr Glu Asp Met Leu Pro Val Met Ala Glu Lys Met Asp Val 35 40 45 Glu Glu Phe Val Ser Glu Leu Cys Lys Gly Phe Ser Leu Leu Ala Asp 50 55 60 Pro Glu Arg His Leu Ile Thr Ala Glu Ser Leu Arg Arg Asn Ser Gly 65 70 75 Ile Leu Gly Ile Glu Gly Met Ser Lys Glu Asp Ala Gln Gly Met Val 80 85 90 95 Arg Glu Gly Asp Leu Asp Gly Asp Gly Ala Leu Asn Gln Thr Glu Phe 100 105 110 Cys Val Leu Met Val Arg Leu Ser Pro Glu Met Met Glu Asp Ala Glu 115 120 125 Thr Trp Leu Glu Lys Ala Leu Thr Gln Glu Leu Cys Asn His Asn Leu 130 135 140 Ser Ser Met Pro 145
20	Asp Thr Ser Tyr Met Glu Ser Ser Asn Leu Tyr Asp Phe Asn His Glu	20	(2) INFORMATION FOR SEQ ID NO:385:
25	Lys Lys Gly Leu Ser Lys Ser Ile Gln Asn Ile Pro Met Phe Tyr	25	(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 135 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide (ix) FEATURE: (A) NAME/KEY: poptide (B) LOCATION: 1..135 (D) OTHER INFORMATION: / Ceres Seq. ID 2026209 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:385: Met Ile Leu Leu Lys Phe Thr Tyr Val Ser Phe Thr Ile Thr Cys 1 5 10 15 Val Ile Phe Ile Leu Ala Ser Ala Ser Val Val Val Ala Ile Val 20 25 30 Gly Ala Ile Glu Gly His Thr Asp Ile Gly Phe Leu Gln Gly Ser 35 40 45 Val Leu Gly Val Val Ala Gly Val Ile Thr Ala Val Gln Leu Phe Gly 50 55 60 Pro Val Leu His Ser Asp Gln Pro Leu Ser Lys Val Ala Leu Leu Arg 65 70 75 80 Arg Val Val Asn Gly Lys Ala Ile Met Gly Leu Val Arg Pro Phe Val 85 90 95 Leu Lys Ala Tyr Gln Trp Gln Ile Ile Ala Leu Asp Thr Ser Tyr Met 100 105 110 Glu Ser Ser Asn Leu Tyr Asp Phe Asn His Glu Lys Lys Gly Leu Ser 115 120 125 Lys Ser Ser Ile Gln Asn Ile Pro Met Phe Tyr Asn Arg Ser Glu His 130 135 140 Gln Thr Lys Ser Ser Cys Ser Ile Cys Leu Gln Asp Trp Glu Gly 145 150 155 160 Glu Val Gly Arg Lys Leu Ala Arg Cys Gly His Thr Phe His Met Asn 165 170 175 Cys Ile Asp Glu Trp Leu Leu Arg Gln Glu Thr Cys Pro Ile 180 185 190 195 200 205
30	Thr Phe His Met Asn Cys Ile Asp Glu Trp Leu Leu Arg Gln Glu Thr	30	(2) INFORMATION FOR SEQ ID NO:386:
35	Cys Pro Ile	35	(1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 206 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: (D) TOPOLOGY: linear (ii) MOLECULE TYPE: peptide (ix) FEATURE: (A) NAME/KEY: poptide (B) LOCATION: 1..206 (D) OTHER INFORMATION: / Ceres Seq. ID 2025547 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:386: Met Ile Leu Leu Lys Phe Thr Tyr Val Ser Phe Thr Ile Thr Cys 1 5 10 15 Val Ile Phe Ile Leu Ala Ser Ala Ser Val Val Val Ala Ile Val 20 25 30 Gly Ala Ile Glu Gly His Thr Asp Ile Gly Phe Leu Gln Gly Ser 35 40 45 Val Leu Gly Val Val Ala Gly Val Ile Thr Ala Val Gln Leu Phe Gly 50 55 60 Pro Val Leu His Ser Asp Gln Pro Leu Ser Lys Val Ala Leu Leu Arg 65 70 75 80 Arg Val Val Asn Gly Lys Ala Ile Met Gly Leu Val Arg Pro Phe Val 85 90 95 Leu Lys Ala Tyr Gln Trp Gln Ile Ile Ala Leu Asp Thr Ser Tyr Met 100 105 110 Glu Ser Ser Asn Leu Tyr Asp Phe Asn His Glu Lys Lys Gly Leu Ser 115 120 125 Lys Ser Ser Ile Gln Asn Ile Pro Met Phe Tyr Asn Arg Ser Glu His 130 135 140 Gln Thr Lys Ser Ser Cys Ser Ile Cys Leu Gln Asp Trp Glu Gly 145 150 155 160 Glu Val Gly Arg Lys Leu Ala Arg Cys Gly His Thr Phe His Met Asn 165 170 175 Cys Ile Asp Glu Trp Leu Leu Arg Gln Glu Thr Cys Pro Ile 180 185 190 195 200 205

515

(A) NAME/KEY: peptide  
 (B) LOCATION: 1..135  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2026209  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:385:  
 5 Met Glu Pro Thr Glu Lys Ser Met Leu Leu Glu Thr Thr Ser Thr Thr  
 1 5 10 15  
 Lys Met Glu Thr Lys Tyr Glu Asp Met Leu Pro Val Met Ala Glu Lys  
 20 25 30  
 Met Asp Val Glu Glu Phe Val Ser Glu Leu Cys Lys Gly Phe Ser Leu  
 35 40 45  
 Leu Ala Asp Pro Glu Arg His Leu Ile Thr Ala Glu Ser Leu Arg Arg  
 50 55 60  
 Asn Ser Gly Ile Leu Gly Ile Glu Gly Met Ser Lys Glu Asp Ala Gln  
 65 70 75  
 Gly Met Val Arg Glu Gly Asp Leu Asp Gly Asp Gly Ala Leu Asn Gln  
 80 85 90  
 Thr Glu Phe Cys Val Leu Met Val Arg Leu Ser Pro Glu Met Met Glu  
 100 105 110  
 Asp Ala Glu Thr Trp Leu Glu Lys Ala Leu Thr Gln Glu Leu Cys Asn  
 115 120 125  
 His Asn Leu Ser Ser Met Pro  
 130  
 (2) INFORMATION FOR SEQ ID NO:386:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 128 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..128  
 (D) OTHER INFORMATION: / Ceres Seq. ID 20..1210  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:386:  
 35 Met Leu Leu Glu Thr Thr Ser Thr Thr Lys Met Glu Thr Lys Tyr Glu  
 1 5 10 15  
 Asp Met Leu Pro Val Met Ala Glu Lys Met Asp Val Glu Glu Phe Val  
 20 25 30  
 Ser Glu Leu Cys Lys Gly Phe Ser Leu Leu Ala Asp Pro Glu Arg His  
 35 40 45  
 Leu Ile Thr Ala Glu Ser Leu Arg Arg Asn Ser Gly Ile Leu Gly Ile  
 50 55 60  
 Glu Gly Met Ser Lys Glu Asp Ala Gln Gly Met Val Arg Glu Gly Asp  
 65 70 75  
 Leu Asp Gly Asp Gly Ala Leu Asn Gln Thr Glu Phe Cys Val Leu Met  
 80 85 90  
 Val Arg Leu Ser Pro Glu Met Met Glu Asp Ala Glu Thr Trp Leu Glu  
 100 105 110  
 Lys Ala Leu Thr Gln Glu Leu Cys Asn His Asn Leu Ser Ser Met Pro  
 115 120 125  
 (2) INFORMATION FOR SEQ ID NO:387:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 552 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..552  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2026982  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:387:

516

atgaagata aaacgtttt tcttaatta tggatgtct cttttatat agtgcacat 60  
 tctcatag cacttacaag tctgacact tatacttct ttcgaaatct cctctatcc 120  
 cttcaaaaa tgaagctctc tggcggttt atcccgctg cttctcttt gtaatgta 180  
 ttcattgcca cagggaagg tccagtcacc gtagagcag gcaactgtg gtaaatgta 240  
 catagttcca aggtccatg tgtgacaca caaactgtg caaactgtg ccaacaaga 300  
 ggtctgcgc gaggtaaatg ccgtggaatc cgtgcgtt gtaactgac aagacatgc 360  
 ggtccacc attccatga cccaactt cgtccatcg taagtggtt actctctct 420  
 taccaaatc ttcgtacgg taccatgag taactgacat gagtggttc tggataagt 480  
 catgttttg tgtttccg gtttaagt atgttaaat caattaagg ctttaatat 540  
 attgatat gg  
 (2) INFORMATION FOR SEQ ID NO:388:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 120 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..120  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2026983  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:388:  
 1 Met Lys Tyr Lys Asn Val Phe Leu Lys Leu Trp Met Ser Ser Leu Tyr  
 5 10 15  
 25 Ile Val Pro Thr Ser Ser Leu Ala Phe Tyr Ser Leu His Thr Tyr His  
 20 25 30  
 Phe Phe Arg Asn Leu Ser Leu Ser Lys Met Lys Leu Ser Val  
 35 40 45  
 Arg Phe Ile Ser Ala Ala Leu Leu Leu Phe Met Val Phe Ile Ala Thr  
 50 55 60  
 Gly Met Gly Pro Val Thr Val Glu Ala Arg Thr Cys Glu Ser Lys Ser  
 65 70 75  
 His Arg Phe Lys Gly Pro Cys Val Ser Thr Thr His Asn Cys Ala Asn Val  
 80 85 90  
 35 Cys His Asn Glu Gly Phe Gly Gly Lys Cys Arg Gly Phe Arg Arg  
 100 105 110  
 Arg Cys Tyr Cys Thr Arg His Cys  
 115 120  
 (2) INFORMATION FOR SEQ ID NO:389:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 109 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..109  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2026984  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:389:  
 1 Met Ser Ser Leu Tyr Ile Val Pro Thr Ser Ser Leu Ala Phe Tyr Ser  
 5 10 15  
 Leu His Thr Tyr His Phe Arg Asn Leu Ser Leu Ser Leu Ser Lys  
 20 25 30  
 55 Met Lys Leu Ser Val Arg Phe Ile Ser Ala Ala Leu Leu Leu Phe Met  
 35 40 45  
 Val Phe Ile Ala Thr Gly Met Gly Pro Val Thr Val Glu Ala Arg Thr  
 50 55 60  
 Cys Glu Ser Lys Ser His Arg Phe Lys Gly Pro Cys Val Ser Thr His  
 65 70 75  
 60 Asn Cys Ala Asn Val Cys His Asn Glu Gly Phe Gly Gly Lys Cys  
 80 85 90

Arg Gly Phe Arg Arg Arg Cys Tyr Cys Thr Arg His Cys  
100 105

(2) INFORMATION FOR SEQ ID NO:390:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 77 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..77

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:390:

Met Lys Leu Ser Val Arg Phe Ile Ser Ala Ala Leu Leu Leu Phe Met  
1 5 10 15

Val Phe Ile Ala Thr Gly Met Gly Pro Val Thr Val Glu Ala Arg Thr  
20 25 30

Cys Glu Ser Lys Ser His Arg Phe Lys Gly Pro Cys Val Ser Thr His  
35 40 45

Asn Cys Ala Asn Val Cys His Asn Glu Gly Phe Gly Gly Cys Cys  
50 55 60

Arg Gly Phe Arg Arg Arg Cys Tyr Cys Thr Arg His Cys  
65 70 75

(2) INFORMATION FOR SEQ ID NO:391:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 66 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..66

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:391:

Met Asp Val Phe Ser Leu Tyr Ser Ala His Phe Leu Ile Ser Ile Leu  
1 5 10 15

Gln Ser Ala His Leu Ser Leu Leu Ser Lys Ser Leu Ser Ile Pro Leu  
20 25 30

Lys Asn Glu Ala Leu Cys Ala Phe Tyr Leu Arg Cys Ser Ser Leu Val  
35 40 45

His Gly Ile His Cys His Arg Asp Gly Ser Ser His Arg Gly Gly Thr  
50 55 60

His Val  
65

(2) INFORMATION FOR SEQ ID NO:392:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1760 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..1760

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:392:

ctcgctttcat gagagaacaa aaaaacttcta tctttctgac aacgggtgag gtgtgtgatg  
120  
ttctcgcga tgcgaagac agtcacaacg atctctctt ctttctctt tcttctctg  
180  
ctttcttcaa atagagcttc ttgaaacga agagctttg tctgtttcg ataccaggg  
240  
ctttacccctc gtgcttaag attgagagctt taatgcttcc ggcacttga tttatcttt

518

ccggcagctc tatgaagta gacagggct tcgcgcttc gacaaagtta cagagtaact  
300  
gaaaagtga tcgagagaga ttgagatcc aagctatttt cagtttccct cctggcttcc  
360  
tgacaagaaa cgttcagctc gagaacacga aacagctcaa acaagactt ctggaagcca  
420  
ttcgccctct tgaagctggc gttacggctt cgcctgatga ccagcttcgg attgatcagt  
480  
tagcgcttaa agtgaagca gttataccca ccaagagacc ttgaaagttc gatttggtca  
540  
atggaaaatg ggagctcatt tatacaacct cgtcttcgat ttgcaggca aagaaacca  
600  
ggctcttaag atcaataacc aacaccaat ctatcaatgt ggatacactt aagtgcaaa  
660  
acaaggagac ttggctcttc tataactcgg taactcgga cataaaacct ctcaactga  
720  
aagaagttgc tgrgaactc caagtgttca cctgagatta cctatctcgg atttattct ataaagcac  
780  
ctgatacggc ccgcggtgaa ctgagatta cctatctcgg agagaaacta cggitatcaa  
840  
gagatcttc attcatgct catgctcgtt tccatgacct tctacatcg ttctcttga  
900  
tttatgaatc tcaataaca ctacaagta accaatag aagtgagttg ttatggctcc  
960  
gttcaaaaga tcaacaacgg agaaagattg aagatggtc aagcagtg ttatggctcc  
1020  
ttggttccgt cgaacacaga aagaagtc ttcagaaac acccagacca ctcttaaaa  
1080  
cttcattgga aaagctttg cctgtttga tacccttgc ttacccttg tctttaaaga  
1140  
ttcgagtttt aagctctcgg ggcgcggat ttaactacc atcgacagt ggtctcgcg  
1200  
ttccgacgaa gtactcagt attcgaagag gtgctcggga aagattgaga atccaagctg  
1260  
tttcaacctt tctccaaga aacgrrggag ccgagaagcg gaacacgctc aaacatgac  
1320  
ttgtcgaagc cattgacct ctcgaactcg gtgcactcg ctcgctgat gatcagcttc  
1380  
tcgagtttt aagctctcgg ggcgcggat ttaactacc atcgacagt ggtctcgcg  
1440  
ttccgacgaa gtactcagt attcgaagag gtgctcggga aagattgaga atccaagctg  
1500  
tcgatttgat caatgggaaa tggagcttca tttaacaac atctgtcgg cctctgaag  
1560  
caagaacac aaggttctta agactgttaa ctaactacca atgatacat atgatacac  
1620  
taaagtgca aagaagggag actgacctt tctaaactc gtaacttga gacttgacac  
1680  
ccctcaactc gaagacggtt gctggaacac ttcaagttt taaattctc ggtttattc  
1740  
aactacgga caaacittga

(2) INFORMATION FOR SEQ ID NO:393:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 527 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: pept-3e

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..527

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:393:

Met Ala Leu Pro Ser Cys Leu Lys Thr Gly Ala Leu Met Ser Pro Ala  
1 5 10 15

Thr Gly Phe Asn Phe Ser Gly Ser Leu Met Lys Ser Asp Ser Gly Phe  
20 25 30

Ala Val Pro Thr Lys Leu Gln Ser Thr Arg Lys Gly Asp Arg Glu Arg  
35 40 45

Leu Arg Val Gln Ala Ile Phe Ser Phe Pro Pro Ala Phe Leu Thr Arg  
50 55 60

Asn Gly Arg Ala Glu Lys Gln Lys Gln Lys Gln Glu Leu Leu Glu  
65 70 75 80

Ala Ile Glu Pro Leu Glu Arg Gly Ala Thr Ala Ser Pro Asp Asp Gln  
85 90 95

Leu Arg Ile Asp Gln Leu Ala Arg Lys Val Glu Ala Val Asn Pro Thr  
100 105 110

Lys Glu Pro Leu Lys Ser Asp Leu Val Asn Gly Lys Trp Glu Leu Ile  
115 120 125

Tyr Thr Thr Ser Ala Ser Ile Leu Gln Ala Lys Lys Pro Arg Phe Leu  
130 135 140

Arg Ser Ile Thr Asn Tyr Gln Ser Ile Asn Val Asp Thr Leu Lys Val  
145 150 155

Gln Asn Met Glu Thr Trp Pro Phe Tyr Asn Ser Val Thr Gly Asp Ile  
165 170 175

Lys Pro Leu Asn Ser Lys Lys Val Ala Val Lys Leu Gln Val Phe Lys  
180 185 190



519  
 Ile Leu Gly Phe Ile Pro Ile Lys Ala Pro Asp Ser Ala Arg Gly Glu  
 195 200 205  
 Leu Glu Ile Thr Tyr Val Asp Glu Glu Leu Arg Leu Ser Arg Asp Leu  
 210 215 220  
 Ser Phe Met Ser His Val Trp Phe His Asp Leu Ser Thr Leu Phe Leu  
 225 230 235 240  
 Leu Ile Tyr Glu Ser Ser Ile Thr Leu Gln Val Lys Pro Ile Arg Ser  
 245 250 255  
 Glu Leu Val Met Val Arg Tyr Lys Asp His Asn Arg Arg Arg Ile Glu  
 260 265 270  
 Glu Trp Ser Lys Arg Trp Ser Glu Leu Val Gly Ser Val Glu Thr Gly  
 275 280 285  
 Arg Lys Ser Leu Ser Glu Thr Thr Gln Pro Leu Phe Lys Thr Pro Leu  
 290 295 300  
 Glu Lys Leu Cys Pro Val Leu Ile Ser Met Ala Leu Pro Trp Cys Leu  
 305 310 315 320  
 Lys Thr Gly Val Leu Thr Ser Pro Ala Ala Gly Phe Asn His Pro Ser  
 325 330 335  
 Asp Ser Gly Phe Ala Val Pro Thr Lys Leu Leu Ser Ile Arg Lys Gly  
 340 345 350  
 Asp Arg Glu Arg Leu Arg Ile Gln Ala Val Phe Ser Phe Pro Pro Arg  
 355 360 365  
 Asn Gly Gly Ala Glu Lys Arg Lys Gln Leu Lys His Glu Leu Val Glu  
 370 375 380  
 Ala Ile Glu Pro Leu Glu Arg Gly Ala Thr Ala Ser Pro Asp Asp Gln  
 385 390 395 400  
 Leu Leu Ile Asp Gln Leu Ala Arg Lys Val Glu Ala Val Asn Pro Thr  
 405 410 415  
 Lys Glu Pro Leu Lys Ser Asp Leu Ile Asn Gly Lys Trp Glu Leu Ile  
 420 425 430  
 Tyr Thr Thr Ser Ala Ala Ile Leu Gln Ala Lys Lys Pro Arg Phe Leu  
 435 440 445  
 Arg Ser Leu Thr Asn Tyr Gln Cys Ile Asn Met Asp Thr Leu Lys Val  
 450 455 460  
 Gln Arg Met Glu Thr Trp Pro Phe Tyr Asn Ser Val Thr Gly Asp Leu  
 465 470 475 480  
 Thr Pro Leu Asn Ser Lys Thr Val Ala Val Lys Leu Gln Val Phe Lys  
 485 490 495  
 Ile Leu Gly Phe Ile Pro Val Lys Ala Pro Asp Gly Thr Ala Arg Gly  
 500 505 510  
 Glu Leu Glu Ile Thr Tyr Val Asp Glu Glu Leu Arg Asp Lys Leu  
 515 520 525  
 (2) INFORMATION FOR SEQ ID NO:394:  
 (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 515 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ix) MOLECULE TYPE: peptide  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..515  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2027302  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:394:  
 Met Ser Pro Ala Thr Gly Phe Asn Phe Ser Gly Ser Leu Met Lys Ser  
 1 5 10 15  
 Asp Ser Gly Phe Ala Val Pro Thr Lys Leu Gln Ser Thr Arg Lys Gly  
 20 25 30  
 Asp Arg Glu Arg Leu Arg Val Gln Ala Ile Phe Ser Phe Pro Pro Ala  
 35 40 45  
 Phe Leu Thr Arg Asn Gly Arg Ala Glu Lys Gln Lys Gln Leu Lys Gln  
 50 55 60

520  
 Glu Leu Leu Glu Ala Ile Glu Pro Leu Glu Arg Gly Ala Thr Ala Ser  
 65 70 75 80  
 Pro Asp Asp Gln Leu Arg Ile Asp Gln Leu Ala Arg Lys Val Glu Ala  
 85 90 95  
 Val Asn Pro Thr Lys Glu Pro Leu Lys Ser Asp Leu Val Asn Gly Lys  
 100 105 110  
 Trp Glu Leu Ile Tyr Thr Thr Ser Ala Ser Ile Leu Gln Ala Lys Lys  
 115 120 125  
 Pro Arg Phe Leu Arg Ser Ile Thr Asn Tyr Gln Ser Ile Asn Val Asp  
 130 135 140  
 Thr Leu Lys Val Gln Asn Met Glu Thr Trp Pro Phe Tyr Asn Ser Val  
 145 150 155 160  
 Thr Gly Asp Ile Lys Pro Leu Asn Ser Lys Lys Val Ala Val Lys Leu  
 165 170 175  
 Gln Val Phe Lys Ile Leu Gly Phe Ile Pro Ile Lys Ala Pro Asp Ser  
 180 185 190  
 Ala Arg Gly Glu Leu Glu Ile Thr Tyr Val Asp Glu Glu Leu Arg Leu  
 195 200 205  
 Ser Arg Asp Leu Ser Phe Met Ser His Val Trp Phe His Asp Leu Ser  
 210 215 220  
 Thr Leu Phe Leu Ile Tyr Glu Ser Ser Ile Thr Leu Gln Val Lys  
 225 230 235 240  
 Pro Ile Arg Ser Glu Leu Val Met Val Arg Tyr Lys Asp His Asn Arg  
 245 250 255  
 Arg Arg Ile Glu Glu Trp Ser Lys Arg Trp Ser Glu Leu Val Gly Ser  
 260 265 270  
 Val Glu Thr Gly Arg Lys Ser Leu Ser Glu Thr Thr Gln Pro Leu Phe  
 275 280 285  
 Lys Thr Pro Leu Glu Lys Leu Cys Pro Val Leu Ile Ser Met Ala Leu  
 290 295 300  
 Pro Trp Cys Leu Lys Thr Gly Val Leu Thr Ser Pro Ala Ala Gly Phe  
 305 310 315 320  
 Asn His Pro Ser Asp Ser Gly Phe Ala Val Pro Thr Lys Leu Leu Ser  
 325 330 335  
 Ile Arg Lys Gly Asp Arg Glu Arg Leu Arg Ile Gln Ala Val Phe Ser  
 340 345 350  
 Phe Pro Pro Arg Asn Gly Ala Glu Lys Arg Lys Gln Leu Lys His  
 355 360 365  
 Glu Leu Val Glu Ala Ile Glu Pro Leu Glu Arg Gly Ala Thr Ala Ser  
 370 375 380  
 Pro Asp Asp Gln Leu Leu Ile Asp Gln Leu Arg Lys Val Glu Ala  
 385 390 395 400  
 Val Asn Pro Thr Lys Glu Pro Leu Lys Ser Asp Leu Ile Asn Gly Lys  
 405 410 415  
 Trp Glu Leu Ile Tyr Thr Thr Ser Ala Ala Ile Leu Gln Ala Lys Lys  
 420 425 430  
 Pro Arg Phe Leu Arg Ser Leu Thr Asn Tyr Gln Cys Ile Asn Met Asp  
 435 440 445  
 Thr Leu Lys Val Gln Arg Met Glu Thr Trp Pro Phe Tyr Asn Ser Val  
 450 455 460  
 Thr Gly Asp Leu Thr Pro Leu Asn Ser Lys Thr Val Ala Val Lys Leu  
 465 470 475 480  
 Gln Val Phe Lys Ile Leu Gly Phe Ile Pro Val Lys Ala Pro Asp Gly  
 485 490 495  
 Thr Ala Arg Gly Glu Leu Glu Ile Thr Tyr Val Asp Glu Glu Leu Arg  
 500 505 510  
 Asp Lys Leu  
 515

(2) INFORMATION FOR SEQ ID NO:395:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 502 amino acids  
 (B) TYPE: amino acid

521

(C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..502  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2027303  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:395:  
 Met Lys Ser Asp Ser Gly Phe Ala Val Pro Thr Lys Leu Gln Ser Thr  
 1 5 10 15  
 Arg Lys Gly Asp Arg Glu Arg Leu Arg Val Gln Ala Ile Phe Ser Phe  
 20 25 30  
 Pro Pro Ala Phe Leu Thr Arg Asn Gly Arg Ala Glu Lys Gln Lys Gln  
 35 40 45  
 Leu Lys Gln Glu Leu Leu Glu Ala Ile Glu Pro Leu Glu Arg Gly Ala  
 50 55 60  
 Thr Ala Ser Pro Asp Asp Gln Leu Arg Ile Asp Gln Leu Ala Arg Lys  
 65 70 75  
 Val Glu Ala Val Asn Pro Thr Lys Glu Pro Leu Lys Ser Asp Leu Val  
 80 85 90  
 Asn Gly Lys Trp Glu Leu Ile Tyr Thr Ser Ala Ser Ile Leu Gln  
 100 105 110  
 Ala Lys Lys Pro Arg Phe Leu Arg Ser Ile Thr Asn Tyr Gln Ser Ile  
 115 120 125  
 Asn Val Asp Thr Leu Lys Val Gln Asn Met Glu Thr Trp Pro Phe Tyr  
 130 135 140  
 Asn Ser Val Thr Gly Asp Ile Lys Pro Leu Asn Ser Lys Lys Val Ala  
 145 150 155  
 Val Lys Leu Gln Val Phe Lys Ile Leu Gly Phe Ile Pro Ile Lys Ala  
 160 165 170  
 Pro Asp Ser Ala Arg Gly Glu Leu Glu Ile Thr Tyr Val Asp Glu Glu  
 175 180 185  
 Leu Arg Leu Ser Arg Asp Leu Ser Phe Met Ser His Val Trp Phe His  
 190 195 200  
 Asp Leu Ser Thr Leu Phe Leu Leu Ile Tyr Glu Ser Ser Ile Thr Leu  
 205 210 215  
 Gln Val Lys Pro Ile Arg Ser Glu Leu Val Met Val Arg Tyr Lys Asp  
 220 225 230  
 His Asn Arg Arg Ile Glu Glu Trp Ser Lys Arg Trp Ser Glu Leu  
 235 240 245  
 Val Gly Ser Val Glu Thr Gly Arg Lys Ser Leu Ser Glu Thr Thr Gln  
 250 255 260  
 Pro Leu Phe Lys Thr Pro Leu Glu Lys Leu Cys Pro Val Leu Ile Ser  
 265 270 275  
 Met Ala Leu Pro Trp Cys Leu Lys Thr Gly Val Leu Thr Ser Pro Ala  
 280 285 290  
 Ala Gly Phe Asn His Pro Ser Asp Ser Gly Phe Ala Val Pro Thr Lys  
 295 300 305  
 Leu Leu Ser Ile Arg Lys Gly Asp Arg Glu Arg Leu Arg Ile Gln Ala  
 310 315 320  
 Val Phe Ser Phe Pro Pro Arg Asn Gly Ala Glu Lys Arg Lys Gln  
 325 330 335  
 Leu Lys His Glu Leu Val Glu Ala Ile Glu Pro Leu Glu Arg Gly Ala  
 340 345 350  
 Thr Ala Ser Pro Asp Asp Gln Leu Leu Ile Asp Gln Leu Ala Arg Lys  
 355 360 365  
 Val Glu Ala Val Asn Pro Thr Lys Glu Pro Leu Lys Ser Asp Leu Ile  
 370 375 380  
 Asn Gly Lys Trp Glu Leu Ile Tyr Thr Ser Ala Ala Ile Leu Gln  
 385 390 395  
 Ala Lys Lys Pro Arg Phe Leu Arg Ser Leu Thr Asn Tyr Gln Cys Ile  
 400 405 410  
 415 420 425 430 440 450

522

Asn Met Asp Thr Leu Lys Val Gln Arg Met Glu Thr Trp Pro Phe Tyr  
 435 440 445  
 Asn Ser Val Thr Gly Asp Leu Thr Pro Leu Asn Ser Lys Thr Val Ala  
 450 455 460  
 Val Lys Leu Gln Val Phe Lys Ile Leu Gly Phe Ile Pro Val Lys Ala  
 465 470 475  
 Pro Asp Gly Thr Ala Arg Gly Glu Leu Glu Ile Thr Tyr Val Asp Glu  
 480 485 490  
 Glu Leu Arg Asp Lys Leu  
 500 505  
 (2) INFORMATION FOR SEQ ID NO:396:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 539 base-pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..539  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2027375  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:396:  
 accgtgttc ataatatgt cgcagagaga agcgatacca agagtcgcgc tgcgtgttt  
 60  
 cattctaac ggaactcaa tcttattagg tcgcgcctgt tctcaatcg gcaactccac  
 120  
 ttctgtctt ccggtggcg acctgaatt cggagagagc ttgagaagt gtgcagcgag  
 180  
 agagaatcg gaggaacag gctaaagat tgaagaatg agcttttga cctgtacaa  
 240  
 caatgtctc aaagaacac caagcctac acatcactc tctgttcca tactgtcgt  
 300  
 gttgtgat ccaagtcaag aaccgaaga tatggaacca gagaagtgt aggtatgga  
 360  
 ttgtcatat tggagaatc taccaaagc ttgttttgg ccaatgaga aattgtttg  
 420  
 aagtggttc aatcttcca ctacgttgg tggagactaa tagatgaag agtcaatgt  
 480  
 tgatttggga ttgaattgt cacaattgg gcaattggtc tagtggtagt atctygttt  
 (2) INFORMATION FOR SEQ ID NO:397:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 152 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..152  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2027376  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:397:  
 Leu Cys Ser Ile Ile Met Ser Thr Gly Glu Ala Ile Pro Arg Val Ala  
 1 5 10  
 Val Val Val Phe Ile Leu Asn Gly Asn Ser Ile Leu Leu Gly Arg Arg  
 15 20 25  
 Arg Ser Ser Ile Gly Asn Ser Thr Phe Ala Leu Pro Gly Gly His Leu  
 30 35 40  
 Glu Phe Gly Glu Ser Phe Glu Glu Cys Ala Ala Arg Glu Val Met Glu  
 45 50 55  
 Glu Thr Gly Leu Lys Ile Glu Lys Met Lys Leu Leu Thr Val Thr Asn  
 60 65 70  
 Asn Val Phe Lys Glu Ala Pro Thr Pro Ser His Tyr Val Ser Val Ser  
 75 80 85  
 Ile Arg Ala Val Leu Val Asp Pro Ser Gln Glu Pro Lys Asn Met Glu  
 90 95 100  
 Pro Glu Lys Cys Glu Gly Trp Asp Trp Tyr Asp Trp Glu Asn Leu Pro  
 105 110 115  
 Lys Pro Leu Phe Trp Pro Leu Glu Lys Leu Phe Gly Ser Gly Phe Asn  
 120 125 130  
 Pro Phe Thr His Gly Gly Asp  
 135 140 145

523

- 145  
(2) INFORMATION FOR SEQ ID NO:398:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 147 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..147  
(D) OTHER INFORMATION: / Ceres Seq. ID 2027377  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:398:  
Met Ser Thr Gly Glu Ala Ile Pro Arg Val Ala Val Val Phe Ile  
1 Met Ser Thr Gly Glu Ala Ile Pro Arg Val Ala Val Val Phe Ile 15  
Leu Asn Gly Asn Ser Ile Leu Leu Gly Arg Arg Ser Ser Ile Gly 30  
Asn Ser Thr Phe Ala Leu Pro Gly Gly His Leu Glu Phe Gly Glu Ser 45  
Phe Glu Glu Cys Ala Ala Arg Glu Val Met Glu Glu Thr Gly Leu Lys 60  
Ile Glu Lys Met Lys Leu Leu Thr Val Thr Asn Asn Val Phe Lys Glu 75  
Ala Pro Thr Pro Ser His Tyr Val Ser Val Ser Ile Arg Ala Val Leu 90  
Val Asp Pro Ser Glu Glu Pro Lys Asn Met Glu Pro Glu Lys Cys Glu 105  
Gly Trp Asp Trp Tyr Asp Trp Glu Asn Leu Pro Lys Pro Leu Phe Trp 120  
Pro Leu Glu Lys Leu Phe Gly Ser Gly Phe Asn Pro Phe Thr His Gly 135  
Gly Gly Asp 145
- 145  
(2) INFORMATION FOR SEQ ID NO:399:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 90 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..90  
(D) OTHER INFORMATION: / Ceres Seq. ID 2027378  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:399:  
Met Glu Glu Thr Gly Leu Lys Ile Glu Lys Met Lys Leu Thr Val 15  
Thr Asn Asn Val Phe Lys Glu Ala Pro Thr Pro Ser His Tyr Val Ser 30  
Val Ser Ile Arg Ala Val Leu Val Asp Pro Ser Gln Glu Pro Lys Asn 45  
Met Glu Pro Glu Lys Cys Glu Gly Trp Asp Trp Tyr Asp Trp Glu Asn 60  
Leu Pro Lys Pro Leu Phe Thr Pro Leu Glu Lys Leu Phe Gly Ser Gly 75  
Phe Asn Pro Phe Thr His Gly Gly Asp 80  
85
- (2) INFORMATION FOR SEQ ID NO:400:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 3983 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single

524

- (D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..3983  
(D) OTHER INFORMATION: / Ceres Seq. ID 2026729  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:400:  
atgaacacga cgcgtttca ctcggtacct cgcgcgtcga ggtaccagcg taagtcgctt 60  
gtcgaagtgt ttgaagtcg taatatctc cctaaagatg gtaacgaag ctcagtcgct 120  
taccgtgttg tgcatttcga tgcctagaag aaacgaacct ccaatagtt cctgacctt 180  
aacctattt ggaacagatg gcttgattc gcgctcccg atcccaaaa ctggtattac 240  
gacgactcg atcagcgtt ttatacagat aaagatttg gtaacgaggt tgcgcggaag 300  
aatatttc tgcgtaggtt taagatctat gaaagccagt tctcgcgaag agtgagaag 360  
ggcttggtt atttccitt ggagaagaag agtcgtctc gctggattcg cggcgagatt 420  
ggatcaaaa tctactatta gcacgaag ccgcacgag ccaacgcggg tggaggtgga 480  
ggagcaaac aacaacaga acagcaaca ttctatcgc cgaacaaga agcagatgaa 540  
caacaacac agcaacaatt tcatctccg ccgcagcaga tgatgaatat accacggag 600  
aaactaatg tagttggtt tgaagaagt aggttttgc atcggttca gctcagcgc 660  
tatcagaaga cacatcagca acctcgggt gttattgtt aagatcac accgagcat 720  
gtaatcgaag gtccaaatga taacctctt caccgaagt ataacctacc tcaacggcca 780  
ccgtctccg cgcacctcc atcggctggg gaagcacatt attaccacc ggaagcgagg 840  
aagatcgaag taggaagacc tccgcgcga gatgaatta ggtttacga gaacacccg 900  
aatgagatt attacctag ggttatcaat agcaaaactg gagagagaga gacgcgatg 960  
gagaagaaga ctcatctcc ttacaacctt gttgagcga tgcagatctt ctcgctcg 1020  
attgtgaag cgcgtggtt accacctaac gagcgcggt atgttaagt accgagctcg 1080  
aacatttgc tcaagttctaa accgcgctt aaccgcctcg gcgaatcgtt tgattcccg 1140  
ggagtgatc aggttttgc tcttgctat aaccgttctg atcctgtgt aactggcg 1200  
actcttga tctctgctg ggaagcttc tgcgagatt ttctcgaggt agttgttt 1260  
gatctctcg aggttcggt tctgaccgc ccggatagtc cgttctctc taagtgttat 1320  
cggtcgaag gctcgcgc ggtacgaac ttctggagaa ttctcggtga catcagctc 1380  
tctattgga ttgtactca ggtatgagtg gcaattcccg aggtctggag ctctgacct 1440  
ccgactgtc ctacacgcg ttctaagtg tctaatcgc cgaacttgc gttctgaga 1500  
gtgacgttc ttgagcaca ggtattacac atagctctc attctccgc gttgactgg 1560  
cttgatctc gtgtgaagc tcaattaggg ttctcagtcg cgtacacag aagagctca 1620  
atgaataac acagtggtc gttctatgg catgagata tgatcttct tctggagag 1680  
ccgttgaag atgtctgtt tctgacgttg gaagaccgga cgaataaaga agcaacct 1740  
ctagacatc ccatgacct agtagctcc atcgagcgc gaattgaga cgtttgttg 1800  
ccgtcgaat gacacctt ggaagagaa ggtgaggtg gaggtagg agtagacct 1860  
ggagtggtg gtgtgtgtg acctattgt ggaagagata gcttaact tctctcga 1920  
ggtggtatc atgtcttga agagcgctg cagtatca cgaattccg tccagcggt 1980  
aagcagtat ggaacccgc gattgaaat ctgagttgg ggaattctg agctgtgg 2040  
ttgttccga tgaagcgaa aaacgaggg aaggttcca ctgacttta ttgttgtt 2100  
aagtcaggga agaatgggt caggactga accatacag acagtttga cccgaggtg 2160  
cacgacagt atactgtga ggtttgat ccttgaccg tctaaactg tgaagcttc 2220  
gacgacagt atactgtga ggtttgat ccttgaccg tctaaactg tgaagcttc 2280  
atagactac taagctctt aggttagca caaagatg cattaaagg ggcgcgacg 2340  
aaatagtag cagcttgct ggtcagcga gaacacatc tggaccaga ggaatttga 2400  
tataattag atcgagatc gcagctatg agcatgaga aaagaaagc gaattgtac 2460  
agaattgtg gtgttttagc ttggcagtg ggtttaga aaaggttga taatcaggg 2520  
atgactaca taagctctt aggttagca caaagatg cattaaagg ggcgcgacg 2580  
aaatagtag cagcttgct ggtcagcga gaacacatc tggaccaga ggaatttga 2640  
tataattag atcgagatc gcagctatg agcatgaga aaagaaagc gaattgtac 2700  
agaattgtg gtgttttagc ttggcagtg ggtttaga aaaggttga taatcaggg 2760  
gggtgagga atccagtag cagctgtgca tctcattc tatctcgtt tctgtttgt 2820  
taccctgatt tggtagtcc gactgatc ctgacttc tgatcagtg gattgtgtac 2880  
taccctgatt tggtagtcc gactgatc ctgacttc tgatcagtg gattgtgtac 2940  
taccctgatt tggtagtcc gactgatc ctgacttc tgatcagtg gattgtgtac 3000  
taccctgatt tggtagtcc gactgatc ctgacttc tgatcagtg gattgtgtac 3060  
taccctgatt tggtagtcc gactgatc ctgacttc tgatcagtg gattgtgtac 3120  
taccctgatt tggtagtcc gactgatc ctgacttc tgatcagtg gattgtgtac 3180  
taccctgatt tggtagtcc gactgatc ctgacttc tgatcagtg gattgtgtac 3240  
taccctgatt tggtagtcc gactgatc ctgacttc tgatcagtg gattgtgtac 3300

525  
ccattctcc ccgcttttc ccgcttttc tccacggaa aatcaccagc aactttctcc 3360  
tgcgagcta tggctccca gaagaggtg aacaaatag atgcacagtg gaagaacaa 3420  
tggacaggag atgattgtt ttcaagagg agtgagcaaa taacgttga tgtttcaag 3480  
aaqctggaga agcgaaagt gtcgagcaac gttgagaat ctggccgct gcaaaagca 3540  
gaaggggttg gactcaact gtatcctt gagaagctta aagttcttc caaagcagg 3600  
gaattgttc tctcagtt ccttgagac ttagtggaa catcgctgc gttcttagc 3660  
tcggctgat tccagctct cagctgctt atgtagcct tgggttgat cccgatgac 3720  
tcaactact taggttgc tagcgctt ttggcggct ccttgagct tacaagggtt 3780  
gtttgttg ttgtttct ttgtttgat ggaactcaag aagctgactg attctttcc 3840  
tgtaaacga acataaccc atgtctgtc caatgatt ttgtcagtt atgattata 3900  
ttataataa taacaacat att 3960

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1276 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..1276  
(D) OTHER INFORMATION: / Ceres Seq. ID 2028730  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:401:  
1 Met Asn Thr Thr Pro Phe His Ser Asp Pro Pro Ser Arg Ile Gln 15  
Arg Lys Leu Val Val Glu Val Val Glu Ala Arg Asn Ile Leu Pro Lys 15  
20 Asp Gly Gln Gly Ser Ser Ser Ala Tyr Val Val Val Asp Phe Asp Ala 30  
Gln Lys Lys Arg Thr Ser Thr Lys Phe Arg Asp Leu Asn Pro Ile Trp 45  
Asn Glu Met Leu Asp Phe Ala Val Ser Asp Pro Lys Asn Met Asp Tyr 60  
65 Asp Glu Leu Asp Ile Glu Val Tyr Asn Asp Lys Arg Phe Gly Asn Gly 75  
Gly Gly Arg Lys Asn His Phe Leu Gly Arg Val Lys Ile Tyr Gly Ser 90  
100 Gln Phe Ser Arg Arg Gly Glu Glu Gly Leu Val Tyr Phe Pro Leu Glu 110  
Lys Lys Ser Val Phe Ser Trp Ile Arg Gly Glu Ile Gly Leu Lys Ile 125  
130 Tyr Tyr Tyr Asp Glu Ala Ala Asp Glu Asp Thr Ala Gly Gly Gly 140  
145 Gly Gln Gln Gln Gln Gln Gln Gln Phe His Pro Pro Gln Gln 155  
Glu Ala Asp Glu Gln Gln His Gln Gln Phe His Pro Pro Gln 170  
Gln Met Met Asn Ile Pro Pro Glu Lys Pro Asn Val Val Val Glu 185  
Glu Gly Arg Val Phe Glu Ser Ala Gln Ser Gln Arg Tyr Thr Glu Thr 195  
His Gln Gln Pro Pro Val Val Ile Val Glu Glu Ser Pro Pro Gln His 205  
225 Val Met Gln Gly Pro Asn Asp Asn His Pro His Arg Asn Asn His 220  
Pro Gln Arg Pro Pro Ser Pro Pro Pro Pro Ser Ala Gly Glu Val 235  
His Tyr Tyr Pro Pro Glu Val Arg Lys Met Gln Val Gly Arg Pro Pro 245  
Gly Gly Asp Arg Ile Arg Val Thr Lys Arg Pro Pro Asn Gly Asp Tyr 255  
260 265 270 280 285 290 300

526  
Ser Pro Arg Val Ile Asn Ser Lys Thr Gly Gly Glu Thr Thr Met 310  
Glu Lys Lys Thr His His Pro Tyr Asn Leu Val Glu Pro Met Gln Tyr 315  
320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800

5 Leu Phe Val Arg Ile Val Lys Ala Arg Gly Leu Pro Pro Asn Glu Ser 340  
Ala Tyr Val Lys Val Arg Thr Ser Asn His Phe Val Arg Ser Lys Pro 345  
350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800

10 Ala Val Asn Arg Pro Gly Glu Ser Val Asp Ser Pro Glu Trp Asn Gln 370  
Val Phe Ala Leu Gly His Asn Arg Ser Asp Ser Ala Val Thr Gly Ala 375  
380 385 390 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800

15 Gly Val Cys Phe Asp Leu Ser Glu Val Pro Val Arg Asp Pro Pro Asp 405  
Thr Leu Glu Ile Ser Ala Trp Asp Ala Ser Ser Glu Ser Phe Leu Gly 410  
415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800

20 Gln Asn Ser Gly Arg Ile Ser Gly Asp Ile Gln Leu Ser Val Trp Ile 440  
Gly Thr Gln Val Asp Glu Ala Phe Pro Glu Ala Trp Ser Ser Asp Ala 445  
450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800

25 Trp Tyr Leu Arg Val Thr Val Leu Glu Ala Gln Asp Leu His Ile Ala 500  
Pro Asn Leu Pro Pro Leu Thr Ala Pro Glu Ile Arg Val Lys Ala Gln 505  
510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800

30 Leu Gly Phe Gln Ser Ala Arg Thr Arg Arg Gly Ser Met Asn Asn His 515  
Ser Gly Ser Phe His Trp His Glu Asp Met Ile Phe Val Ala Gly Glu 520  
525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800

35 Glu Ala Thr Leu Leu Gly His Ala Met Ile Pro Val Ser Ser Ile Glu 580  
Gln Arg Ile Asp Glu Arg Phe Val Pro Ser Lys Trp His Thr Leu Glu 585  
590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800

40 Gly Glu Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly 610  
Gly Gly Gly Pro Tyr Cys Gly Arg Ile Ser Leu Arg Leu Cys Leu Glu 615  
620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800

45 Arg Pro Thr Ala Lys Gln Leu Trp Lys Pro Pro Ile Gly Ile Leu Glu 645  
Leu Gly Ile Leu Gly Ala Arg Gly Leu Leu Pro Met Lys Ala Lys Asn 650  
655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800

50 Gly Gly Lys Gly Ser Thr Asp Ala Tyr Cys Val Ala Lys Tyr Gly Lys 675  
Lys Trp Val Arg Thr Arg Thr Ile Thr Asp Ser Phe Asp Pro Arg Trp 680  
705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800

55 Val Gly Val Phe Asp Asn Trp Arg Met Phe Ser Asp Ala Ser Asp Asp 740  
Arg Pro Asp Thr Arg Ile Gly Lys Ile Arg Ile Arg Val Ser Thr Leu 745  
750 755 760 765 770 775 780 785 790 795 800

60 Glu Ser Asn Lys Val Tyr Thr Asn Ser Tyr Pro Leu Leu Val Leu Leu 755  
Pro Ser Gly Met Lys Lys Met Gly Glu Ile Glu Val Ala Val Arg Phe 760  
765 770 775 780 785 790 795 800

5	Ala Cys Pro Ser Leu Leu Pro Asp Val Cys Ala Ala Tyr Gly Gln Pro	527	(A) LENGTH: 1210 amino acids	528
	Leu Leu Pro Arg Met His Tyr Ile Arg Pro Leu Gly Val Ala Gln Gln	810	(B) TYPE: amino acid	
10	Asp Ala Leu Arg Gly Ala Ala Thr Lys Met Val Ala Ala Trp Leu Ala	820	(C) STRANDEDNESS:	
	Arg Ala Glu Pro Pro Leu Gly Pro Glu Val Val Arg Tyr Met Leu Asp	825	(D) TOPOLOGY: linear	
15	Ala Asp Ser His Ala Trp Ser Met Arg Lys Ser Lys Ala Asn Trp Tyr	830	(ii) MOLECULE TYPE: peptide	
	Arg Ile Val Gly Val Leu Ala Trp Ala Val Gly Leu Ala Lys Trp Leu	835	(ix) FEATURE:	
20	Asp Asn Ile Arg Arg Trp Arg Asn Pro Val Thr Thr Val Leu Val His	840	(A) NAME/KEY: peptide	
	Ile Leu Tyr Leu Val Leu Val Trp Tyr Pro Asp Leu Val Val Pro Thr	845	(B) LOCATION: 1..1210	
25	Ala Phe Leu Tyr Val Val Met Ile Gly Val Trp Tyr Arg Phe Arg	850	(D) OTHER INFORMATION: / Ceres Seq. ID 2028731	
	Pro Lys Ile Pro Ala Gly Met Asp Ile Arg Leu Ser Gln Ala Glu Thr	855	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:402:	
30	Val Asp Pro Asp Glu Leu Asp Glu Phe Asp Thr Ile Pro Ser Ser	860	Met Leu Asp Phe Ala Val Ser Asp Pro Lys Asn Met Asp Tyr Asp Glu	1
	Arg Arg Pro Glu Val Ile Arg Ala Arg Tyr Asp Arg Leu Arg Ile Leu	865	Leu Asp Ile Glu Val Tyr Asn Asp Lys Arg Phe Gly Asn Gly Gly Gly	10
35	Ala Val Arg Val Gln Thr Ile Leu Gly Asp Phe Ala Ala Gln Gly Glu	870	Arg Lys Asn His Phe Leu Gly Arg Val Lys Ile Tyr Gly Ser Gln Phe	20
	Arg Ile Gln Ala Leu Val Ser Trp Arg Asp Pro Arg Ala Thr Lys Leu	875	Ser Arg Arg Gly Glu Glu Gly Leu Val Tyr Phe Pro Leu Glu Lys Lys	30
40	Phe Ile Ala Ile Cys Leu Val Ile Thr Ile Val Leu Tyr Ala Val Pro	880	Ser Val Phe Ser Trp Ile Arg Gly Glu Ile Gly Leu Lys Ile Tyr Tyr	40
	Ala Lys Met Val Ala Val Ala Leu Gly Val Ser Asp Ser Val Pro Thr	885	Tyr Asp Glu Ala Ala Asp Glu Asp Thr Ala Gly Gly Gly Gly Gln	50
45	Ala Lys Gln Asp Thr Lys Glu Ser Leu Lys Lys Ser Phe Ser Ser Leu	890	Gln Gln Gln Gln Gln Gln Phe His Pro Pro Gln Gln Glu Ala	60
	Arg Phe Asp Phe Ser Ser Met Ala Val Val Gly Ala Pro Ile Ser Ser	895	Asp Glu Gln Gln His Gln Gln Phe His Pro Pro Gln Gln Met	70
50	Pro Ala Ala Gln Leu Thr Gln Phe Leu Ser Asn Pro Ile Leu Pro	900	Met Asn Ile Pro Pro Glu Lys Pro Asn Val Val Val Glu Glu Gly	80
	Arg Phe Arg Arg Ser Phe Ser Thr Gly Lys Ser Pro Ala Thr Phe Ser	905	Arg Val Phe Glu Ser Ala Gln Ser Gln Arg Tyr Thr Glu Thr His Gln	90
55	Val Val Ala Met Ala Pro Gln Lys Lys Val Asn Lys Tyr Asp Ala Lys	910	Gln Pro Pro Val Val Ile Val Glu Glu Ser Pro Pro Gln His Val Met	100
	Trp Lys Lys Gln Trp Tyr Gly Ala Gly Leu Phe Phe Glu Gly Ser Glu	915	Gln Gly Pro Asn Asp Asn His Pro His Arg Asn Asp Asn His Pro Gln	110
60	Gln Ile Asn Val Asp Val Phe Lys Lys Leu Glu Lys Arg Lys Val Leu	920	Arg Pro Pro Ser Pro Pro Pro Pro Ser Ala Gly Glu Val His Tyr	120
	Ser Asn Val Glu Lys Ser Gly Leu Leu Ser Lys Ala Glu Gly Leu Gly	925	Tyr Pro Pro Glu Val Arg Lys Met Gln Val Gly Arg Pro Pro Gly Gly	130
65	Leu Thr Leu Ser Ser Leu Glu Lys Leu Lys Val Phe Ser Lys Ala Glu	930	Asp Arg Ile Arg Val Thr Lys Arg Pro Pro Asn Gly Asp Tyr Ser Pro	140
	Asp Leu Gly Leu Ser Leu Leu Glu Asn Leu Ala Gly Thr Ser Pro	935	Arg Val Ile Asn Ser Lys Thr Gly Gly Glu Thr Thr Met Glu Lys	150
70	Ala Val Leu Ala Gly Ala Leu Ala Leu Thr Gly Val Leu Leu Val	940	Lys Thr His His Pro Tyr Asn Leu Val Glu Pro Met Gln Tyr Leu Phe	160
	Ala Val Val Leu Ile Pro Asp Ser Thr Thr Leu Val Val Ala Gln	945	Val Arg Ile Val Lys Ala Arg Gly Leu Pro Pro Asn Glu Ser Ala Tyr	170
75	Ala Val Val Leu Ile Pro Asp Ser Thr Thr Leu Val Val Ala Gln	950	Val Lys Val Arg Thr Ser Asn His Phe Val Arg Ser Lys Pro Ala Val	180
	Ala Val Leu Ala Gly Ala Leu Ala Leu Thr Gly Val Val Leu Val	955	Asn Arg Pro Gly Glu Ser Val Asp Ser Pro Glu Trp Asn Gln Val Phe	190
80	Gly Ser Val Val Leu Asp Gly Leu Gln Glu Ala Asp	960	Asn Arg Pro Gly Glu Ser Val Asp Ser Pro Glu Trp Asn Gln Val Phe	200
	1265	1270	Ala Leu Gly His Asn Arg Ser Asp Ser Ala Val Thr Gly Ala Thr Leu	210
85	1275	1280	Glu Ile Ser Ala Trp Asp Ala Ser Ser Glu Ser Phe Leu Gly Val	220
	1285	1290	Cys Phe Asp Leu Ser Glu Val Pro Val Arg Asp Pro Pro Asp Ser Pro	230
90	1295	1300	Leu Ala Pro Gln Trp Tyr Arg Leu Glu Gly Ser Gly Ala Asp Gln Asn	240
	1305	1310	Ser Gly Arg Ile Ser Gly Asp Ile Gln Leu Ser Val Trp Ile Gly Thr	250
95	1315	1320	385	390
	1325	1330	Gln Val Asp Glu Ala Phe Pro Glu Ala Trp Ser Ser Asp Ala Pro His	400
100	1335	1340	405	410
	1345	1350		

(2) INFORMATION FOR SEQ ID NO:402:  
(i) SEQUENCE CHARACTERISTICS:

5	Val Ala His Thr Arg Ser Lys Val Tyr Gln Ser Pro Lys Leu Trp Tyr	420	425	430	435	440	445	450	455	460	465	470	475	480	485	490	495	500	505	510	515	520	525	530	535	540	545	550	555	560	565	570	575	580	585	590	595	600	605	610	615	620	625	630	635	640	645	650	655	660	665	670	675	680	685	690	695	700	705	710	715	720	725	730	735	740	745	750	755	760	765	770	775	780	785	790	795	800	805	810	815	820	825	830	835	840	845	850	855	860	865	870	875	880	885	890	895	900	905	910	915	920	925	930	935	940	945	950	955	960	965	970	975	980	985	990	995	1000	1005	1010	1015	1020	1025	1030	1035	1040	1045	1050	1055	1060	1065	1070	1075	1080	1085	1090	1095	1100	1105	1110	1115	1120	1125	1130	1135	1140	1145	1150	1155	1160	1165	1170	1175	1180	1185	1190	1195	1200	1205	1210	1215	1220	1225	1230	1235	1240	1245	1250	1255	1260	1265	1270	1275	1280	1285	1290	1295	1300	1305	1310	1315	1320	1325	1330	1335	1340	1345	1350	1355	1360	1365	1370	1375	1380	1385	1390	1395	1400	1405	1410	1415	1420	1425	1430	1435	1440	1445	1450	1455	1460	1465	1470	1475	1480	1485	1490	1495	1500	1505	1510	1515	1520	1525	1530	1535	1540	1545	1550	1555	1560	1565	1570	1575	1580	1585	1590	1595	1600	1605	1610	1615	1620	1625	1630	1635	1640	1645	1650	1655	1660	1665	1670	1675	1680	1685	1690	1695	1700	1705	1710	1715	1720	1725	1730	1735	1740	1745	1750	1755	1760	1765	1770	1775	1780	1785	1790	1795	1800	1805	1810	1815	1820	1825	1830	1835	1840	1845	1850	1855	1860	1865	1870	1875	1880	1885	1890	1895	1900	1905	1910	1915	1920	1925	1930	1935	1940	1945	1950	1955	1960	1965	1970	1975	1980	1985	1990	1995	2000	2005	2010	2015	2020	2025	2030	2035	2040	2045	2050	2055	2060	2065	2070	2075	2080	2085	2090	2095	2100	2105	2110	2115	2120	2125	2130	2135	2140	2145	2150	2155	2160	2165	2170	2175	2180	2185	2190	2195	2200	2205	2210	2215	2220	2225	2230	2235	2240	2245	2250	2255	2260	2265	2270	2275	2280	2285	2290	2295	2300	2305	2310	2315	2320	2325	2330	2335	2340	2345	2350	2355	2360	2365	2370	2375	2380	2385	2390	2395	2400	2405	2410	2415	2420	2425	2430	2435	2440	2445	2450	2455	2460	2465	2470	2475	2480	2485	2490	2495	2500	2505	2510	2515	2520	2525	2530	2535	2540	2545	2550	2555	2560	2565	2570	2575	2580	2585	2590	2595	2600	2605	2610	2615	2620	2625	2630	2635	2640	2645	2650	2655	2660	2665	2670	2675	2680	2685	2690	2695	2700	2705	2710	2715	2720	2725	2730	2735	2740	2745	2750	2755	2760	2765	2770	2775	2780	2785	2790	2795	2800	2805	2810	2815	2820	2825	2830	2835	2840	2845	2850	2855	2860	2865	2870	2875	2880	2885	2890	2895	2900	2905	2910	2915	2920	2925	2930	2935	2940	2945	2950	2955	2960	2965	2970	2975	2980	2985	2990	2995	3000	3005	3010	3015	3020	3025	3030	3035	3040	3045	3050	3055	3060	3065	3070	3075	3080	3085	3090	3095	3100	3105	3110	3115	3120	3125	3130	3135	3140	3145	3150	3155	3160	3165	3170	3175	3180	3185	3190	3195	3200	3205	3210	3215	3220	3225	3230	3235	3240	3245	3250	3255	3260	3265	3270	3275	3280	3285	3290	3295	3300	3305	3310	3315	3320	3325	3330	3335	3340	3345	3350	3355	3360	3365	3370	3375	3380	3385	3390	3395	3400	3405	3410	3415	3420	3425	3430	3435	3440	3445	3450	3455	3460	3465	3470	3475	3480	3485	3490	3495	3500	3505	3510	3515	3520	3525	3530	3535	3540	3545	3550	3555	3560	3565	3570	3575	3580	3585	3590	3595	3600	3605	3610	3615	3620	3625	3630	3635	3640	3645	3650	3655	3660	3665	3670	3675	3680	3685	3690	3695	3700	3705	3710	3715	3720	3725	3730	3735	3740	3745	3750	3755	3760	3765	3770	3775	3780	3785	3790	3795	3800	3805	3810	3815	3820	3825	3830	3835	3840	3845	3850	3855	3860	3865	3870	3875	3880	3885	3890	3895	3900	3905	3910	3915	3920	3925	3930	3935	3940	3945	3950	3955	3960	3965	3970	3975	3980	3985	3990	3995	4000	4005	4010	4015	4020	4025	4030	4035	4040	4045	4050	4055	4060	4065	4070	4075	4080	4085	4090	4095	4100	4105	4110	4115	4120	4125	4130	4135	4140	4145	4150	4155	4160	4165	4170	4175	4180	4185	4190	4195	4200	4205	4210	4215	4220	4225	4230	4235	4240	4245	4250	4255	4260	4265	4270	4275	4280	4285	4290	4295	4300	4305	4310	4315	4320	4325	4330	4335	4340	4345	4350	4355	4360	4365	4370	4375	4380	4385	4390	4395	4400	4405	4410	4415	4420	4425	4430	4435	4440	4445	4450	4455	4460	4465	4470	4475	4480	4485	4490	4495	4500	4505	4510	4515	4520	4525	4530	4535	4540	4545	4550	4555	4560	4565	4570	4575	4580	4585	4590	4595	4600	4605	4610	4615	4620	4625	4630	4635	4640	4645	4650	4655	4660	4665	4670	4675	4680	4685	4690	4695	4700	4705	4710	4715	4720	4725	4730	4735	4740	4745	4750	4755	4760	4765	4770	4775	4780	4785	4790	4795	4800	4805	4810	4815	4820	4825	4830	4835	4840	4845	4850	4855	4860	4865	4870	4875	4880	4885	4890	4895	4900	4905	4910	4915	4920	4925	4930	4935	4940	4945	4950	4955	4960	4965	4970	4975	4980	4985	4990	4995	5000	5005	5010	5015	5020	5025	5030	5035	5040	5045	5050	5055	5060	5065	5070	5075	5080	5085	5090	5095	5100	5105	5110	5115	5120	5125	5130	5135	5140	5145	5150	5155	5160	5165	5170	5175	5180	5185	5190	5195	5200	5205	5210	5215	5220	5225	5230	5235	5240	5245	5250	5255	5260	5265	5270	5275	5280	5285	5290	5295	5300	5305	5310	5315	5320	5325	5330	5335	5340	5345	5350	5355	5360	5365	5370	5375	5380	5385	5390	5395	5400	5405	5410	5415	5420	5425	5430	5435	5440	5445	5450	5455	5460	5465	5470	5475	5480	5485	5490	5495	5500	5505	5510	5515	5520	5525	5530	5535	5540	5545	5550	5555	5560	5565	5570	5575	5580	5585	5590	5595	5600	5605	5610	5615	5620	5625	5630	5635	5640	5645	5650	5655	5660	5665	5670	5675	5680	5685	5690	5695	5700	5705	5710	5715	5720	5725	5730	5735	5740	5745	5750	5755	5760	5765	5770	5775	5780	5785	5790	5795	5800	5805	5810	5815	5820	5825	5830	5835	5840	5845	5850	5855	5860	5865	5870	5875	5880	5885	5890	5895	5900	5905	5910	5915	5920	5925	5930	5935	5940	5945	5950	5955	5960	5965	5970	5975	5980	5985	5990	5995	6000	6005	6010	6015	6020	6025	6030	6035	6040	6045	6050	6055	6060	6065	6070	6075	6080	6085	6090	6095	6100	6105	6110	6115	6120	6125	6130	6135	6140	6145	6150	6155	6160	6165	6170	6175	6180	6185	6190	6195	6200	6205	6210	6215	6220	6225	6230	6235	6240	6245	6250	6255	6260	6265	6270	6275	6280	6285	6290	6295	6300	6305	6310	6315	6320	6325	6330	6335	6340	6345	6350	6355	6360	6365	6370	6375	6380	6385	6390	6395	6400	6405	6410	6415	6420	6425	6430	6435	6440	6445	6450	6455	6460	6465	6470	6475	6480	6485	6490	6495	6500	6505	6510	6515	6520	6525	6530	6535	6540	6545	6550	6555	6560	6565	6570	6575	6580	6585	6590	6595	6600	6605	6610	6615	6620	6625	6630	6635	6640	6645	6650	6655	6660	6665	6670	6675	6680	6685	6690	6695	6700	6705	6710	6715	6720	6725	6730	6735	6740	6745	6750	6755	6760	6765	6770	6775	6780	6785	6790	6795	6800	6805	6810	6815	6820	6825	6830	6835	6840	6845	6850	6855	6860	6865	6870	6875	6880	6885	6890	6895	6900	6905	6910	6915	6920	6925	6930	6935	6940	6945	6950	6955	6960	6965	6970	6975	6980	6985	6990	6995	7000	7005	7010	7015	7020	7025	7030	7035	7040	7045	7050	7055	7060	7065	7070	7075	7080	7085	7090	7095	7100	7105	7110	7115	7120	7125	7130	7135	7140	7145	7150	7155	7160	7165	7170	7175	7180	7185	7190	7195	7200	7205	7210	7215	7220	7225	7
---	---	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	---

531

	Pro	Gln	Gln	Glu	Ala	Asp	Glu	Gln	Gln	His	Gln	Gln	Gln	Phe	His	Pro
				100					105					110		
	Pro	Pro	Gln	Gln	Met	Met	Asn	Ile	Pro	Pro	Glu	Lys	Pro	Asn	Val	Val
			115					120					125			
5	Val	Val	Glu	Glu	Gly	Arg	Val	Phe	Glu	Ser	Ala	Gln	Ser	Gln	Arg	Tyr
			130				135					140				
	Thr	Glu	Thr	His	Gln	Gln	Pro	Pro	Val	Val	Ile	Val	Glu	Glu	Ser	Pro
						150					155					160
10	Pro	Gln	His	Val	Met	Gln	Gly	Pro	Asn	Asp	Asn	His	Pro	His	Arg	Asn
					165					170						175
	Asp	Asn	His	Pro	Gln	Arg	Pro	Pro	Ser	Pro	Pro	Pro	Pro	Pro	Ser	Ala
				180					185						190	
	Gly	Glu	Val	His	Tyr	Tyr	Pro	Pro	Glu	Val	Arg	Lys	Met	Gln	Val	Gly
			195				200						205			
15	Arg	Pro	Pro	Gly	Gly	Asp	Arg	Ile	Arg	Val	Thr	Lys	Arg	Pro	Pro	Asn
				210			215					220				
	Gly	Asp	Tyr	Ser	Pro	Arg	Val	Ile	Asn	Ser	Lys	Thr	Gly	Gly	Gly	Glu
						230					235					240
20	Thr	Thr	Met	Glu	Lys	Lys	Thr	His	His	Pro	Tyr	Asn	Leu	Val	Glu	Pro
					245					250					255	
	Met	Gln	Tyr	Leu	Phe	Val	Arg	Ile	Val	Lys	Ala	Arg	Gly	Leu	Pro	Pro
				260					265					270		
	Asn	Glu	Ser	Ala	Tyr	Val	Lys	Val	Arg	Thr	Ser	Asn	His	Phe	Val	Arg
				275			280						285			
25	Ser	Lys	Pro	Ala	Val	Asn	Arg	Pro	Gly	Glu	Ser	Val	Asp	Ser	Pro	Glu
				290			295					300				
	Trp	Asn	Gln	Val	Phe	Ala	Leu	Gly	His	Asn	Arg	Ser	Asp	Ser	Ala	Val
						310					315					320
30	Thr	Gly	Ala	Thr	Leu	Glu	Ile	Ser	Ala	Trp	Asp	Ala	Ser	Ser	Glu	Ser
					325					330					335	
	Phe	Leu	Gly	Gly	Val	Cys	Phe	Asp	Leu	Ser	Glu	Val	Pro	Val	Arg	Asp
				340					345					350		
	Pro	Pro	Asp	Ser	Pro	Leu	Ala	Pro	Gln	Trp	Tyr	Arg	Leu	Glu	Gly	Ser
				355			360						365			
35	Gly	Ala	Asp	Gln	Asn	Ser	Gly	Arg	Ile	Ser	Gly	Asp	Ile	Gln	Leu	Ser
				370			375					380				
	Val	Trp	Ile	Gly	Thr	Gln	Val	Asp	Glu	Ala	Phe	Pro	Glu	Ala	Trp	Ser
						390					395					400
40	Ser	Asp	Ala	Pro	His	Val	Ala	His	Thr	Arg	Ser	Lys	Val	Tyr	Gln	Ser
					405					410					415	
	Pro	Lys	Leu	Trp	Tyr	Leu	Arg	Val	Thr	Val	Leu	Glu	Ala	Gln	Asp	Leu
				420					425					430		
	His	Ile	Ala	Pro	Asn	Leu	Pro	Pro	Leu	Thr	Ala	Pro	Glu	Ile	Arg	Val
				435			440						445			
45	Lys	Ala	Gln	Leu	Gly	Phe	Gln	Ser	Ala	Arg	Thr	Arg	Arg	Gly	Ser	Met
				450			455					460				
	Asn	Asn	His	Ser	Gly	Ser	Phe	His	Trp	His	Glu	Asp	Met	Ile	Phe	Val
						470					475					480
50	Ala	Gly	Glu	Pro	Leu	Glu	Asp	Cys	Leu	Val	Leu	Met	Val	Glu	Asp	Arg
					485					490					495	
	Thr	Thr	Lys	Glu	Ala	Thr	Leu	Leu	Gly	His	Ala	Met	Ile	Pro	Val	Ser
				500					505					510		
	Ser	Ile	Glu	Gln	Arg	Ile	Asp	Glu	Arg	Phe	Val	Pro	Ser	Lys	Trp	His
				515				520					525			
55	Thr	Leu	Glu	Gly	Glu	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Pro	Gly
				530			535						540			
	Gly	Gly	Gly	Gly	Gly	Gly	Pro	Tyr	Cys	Gly	Arg	Ile	Ser	Leu	Arg	Leu
						550					555					560
60	Cys	Leu	Glu	Gly	Gly	Tyr	His	Val	Leu	Glu	Ala	Ala	His	Val	Cys	
					565					570				575		
	Ser	Asp	Phe	Arg	Pro	Thr	Ala	Lys	Gln	Leu	Trp	Lys	Pro	Pro	Ile	Gly
				580					585					590		

532

	Ile	Leu	Glu	Leu	Gly	Ile	Leu	Gly	Ala	Arg	Gly	Leu	Leu	Pro	Met	Lys
		595						600					605			
	Ala	Lys	Asn	Gly	Gly	Lys	Gly	Ser	Thr	Asp	Ala	Tyr	Cys	Val	Ala	Lys
		610					615						620			
5	Tyr	Gly	Lys	Lys	Trp	Val	Arg	Thr	Arg	Thr	Ile	Thr	Asp	Ser	Phe	Asp
		625				630					635					640
	Pro	Arg	Trp	His	Glu	Gln	Tyr	Thr	Trp	Gln	Val	Tyr	Asp	Pro	Cys	Thr
					645					650					655	
10	Val	Leu	Thr	Val	Gly	Val	Phe	Asp	Asn	Trp	Arg	Met	Phe	Ser	Asp	Ala
				660					665						670	
	Ser	Asp	Asp	Arg	Pro	Asp	Thr	Arg	Ile	Gly	Lys	Ile	Arg	Ile	Arg	Val
			675					680					685			
	Ser	Thr	Leu	Glu	Ser	Asn	Lys	Val	Tyr	Thr	Asn	Ser	Tyr	Pro	Leu	Leu
		690				695						700				
15	Val	Leu	Leu	Pro	Ser	Gly	Met	Lys	Lys	Met	Gly	Glu	Ile	Glu	Val	Ala
		705				710					715					720
	Val	Arg	Phe	Ala	Cys	Pro	Ser	Leu	Leu	Pro	Asp	Val	Cys	Ala	Ala	Tyr
				725						730					735	
20	Gly	Gln	Pro	Leu	Leu	Pro	Arg	Met	His	Tyr	Ile	Arg	Pro	Leu	Gly	Val
				740					745						750	
	Ala	Gln	Gln	Asp	Ala	Leu	Arg	Gly	Ala	Ala	Thr	Lys	Met	Val	Ala	Ala
			755					760					765			
	Trp	Leu	Ala	Arg	Ala	Glu	Pro	Pro	Leu	Gly	Pro	Glu	Val	Val	Arg	Tyr
		770				775						780				
25	Met	Leu	Asp	Ala	Asp	Ser	His	Ala	Trp	Ser	Met	Arg	Lys	Ser	Lys	Ala
		785				790					795					800
	Asn	Trp	Tyr	Arg	Ile	Val	Gly	Val	Leu	Ala	Trp	Ala	Val	Gly	Leu	Ala
				805					810						815	
30	Lys	Trp	Leu	Asp	Asn	Ile	Arg	Arg	Trp	Arg	Asn	Pro	Val	Thr	Thr	Val
				820					825					830		
	Leu	Val	His	Ile	Leu	Tyr	Leu	Val	Leu	Val	Trp	Tyr	Pro	Asp	Leu	Val
			835					840					845			
	Val	Pro	Thr	Ala	Phe	Leu	Tyr	Val	Val	Met	Ile	Gly	Val	Trp	Tyr	Tyr
		850				855					860					
35	Arg	Phe	Arg	Pro	Lys	Ile	Pro	Ala	Gly	Met	Asp	Ile	Arg	Leu	Ser	Gln
		865				870					875					880
	Ala	Glu	Thr	Val	Asp	Pro	Asp	Glu	Leu	Asp	Glu	Glu	Phe	Asp	Thr	Ile
				885					890						895	
40	Pro	Ser	Ser	Arg	Arg	Pro	Glu	Val	Ile	Arg	Ala	Arg	Tyr	Asp	Arg	Leu
				900					905					910		
	Arg	Ile	Leu	Ala	Val	Arg	Val	Gln	Thr	Ile	Leu	Gly	Asp	Phe	Ala	Ala
			915					920					925			
	Gln	Gly	Glu	Arg	Ile	Gln	Ala	Leu	Val	Ser	Trp	Arg	Asp	Pro	Arg	Ala
		930				935						940				
45	Thr	Lys	Leu	Phe	Ile	Ala	Ile	Cys	Leu	Val	Ile	Thr	Ile	Val	Leu	Tyr
		945				950					955					960
	Ala	Val	Pro	Ala	Lys	Met	Val	Ala	Val	Ala	Leu	Gly	Val	Ser	Asp	Ser
				965						970					975	
50	Val	Pro	Thr	Ala	Lys	Gln	Asp	Thr	Lys	Glu	Ser	Leu	Lys	Lys	Ser	Phe
				980					985					990		
	Ser	Ser	Leu	Arg	Phe	Asp	Phe	Ser	Ser	Met	Ala	Val	Val	Gly	Ala	Pro
			995					1000					1005			
	Ile	Ser	Ser	Pro	Ala	Ala	Gln	Leu	Gln	Thr	Gln	Phe	Leu	Ser	Asn	Pro
		1010				1015						1020				
55	Ile	Leu	Pro	Arg	Phe	Arg	Arg	Ser	Phe	Ser	Thr	Gly	Lys	Ser	Pro	Ala
		1025				1030					1035					1040
	Thr	Phe	Ser	Val	Val	Ala	Met	Ala	Pro	Gln	Lys	Lys	Val	Asn	Lys	Tyr
				1045						1050					1055	
	Asp	Ala	Lys	Trp	Lys	Lys	Gln	Trp	Tyr	Gly	Ala	Gly	Leu	Phe	Phe	Glu
			1060						1065					1070		
60	Gly	Ser	Glu	Gln	Ile	Asn	Val	Asp	Val	Phe	Lys	Lys	Leu	Glu	Lys	Arg
			1075					1080						1085		



533

Lys Val Leu Ser Asn Val Glu Lys Ser Gly Leu Leu Ser Lys Ala Glu  
 1090 1095 1100  
 Gly Leu Gly Leu Thr Leu Ser Ser Leu Glu Lys Leu Lys Val Phe Ser  
 1105 1110 1115 1120  
 5 Lys Ala Glu Asp Leu Gly Leu Leu Ser Leu Leu Glu Asn Leu Ala Gly  
 1125 1130 1135  
 Thr Ser Pro Ala Val Leu Ala Ser Ala Ala Leu Pro Ala Leu Thr Ala  
 1140 1145 1150  
 10 Ala Ile Val Ala Val Val Leu Ile Pro Asp Asp Ser Thr Thr Leu Val  
 1155 1160 1165  
 Val Ala Gln Ala Val Leu Ala Gly Ala Leu Ala Leu Thr Gly Val Val  
 1170 1175 1180  
 Leu Leu Val Gly Ser Val Val Leu Asp Gly Leu Gln Glu Ala Asp  
 1185 1190 1195  
 15 (2) INFORMATION FOR SEQ ID NO:404:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 528 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 20 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..528  
 25 (D) OTHER INFORMATION: / Ceres Seq. ID 2029079  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:404:  
 aaaaaactca ctttcacttg cacaaagata aggaaacccat gtctgtgtca gcgatctttg 60  
 gtaccggaat cgtcaccgtc gctgcttctc cggttctccg ccaatttcaa gttccaaaat 120  
 tgggtaaatgg aggtggatta gggatggtga ttgagtgttc gtcgaggcca cagaagaaat 180  
 30 cgacagctca tcacaggaag acgaggccga agaagactca gccttgggac attaagagaa 240  
 agcctactgt gtatgctcct cttcctcctc ttccggcgga atggagtccg tttactcttg 300  
 cttccaacga cgttggtgcc gccgttgctg cttctcctgc cggagatttg gtttcaggct 360  
 ctgcctagtt atatgagcta tctggtggtt tggttgtaat ctagtttgag agcttttggt 420  
 ggtgtttggt tctgttaatg tttttcagtt ttctgctgtc tgtttcttat tatccggttt 480  
 35 atgtttggct ggtgaatgaa tgtcaatgga ccaaactgca attttttt  
 (2) INFORMATION FOR SEQ ID NO:405:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 121 amino acids  
 (B) TYPE: amino acid  
 40 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 45 (B) LOCATION: 1..121  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2029080  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:405:  
 Lys Thr His Phe His Leu His Lys Asp Lys Glu Thr Met Ser Val Ser  
 1 5 10 15  
 50 Ala Ile Phe Gly Thr Gly Ile Val Thr Val Ala Ala Ser Pro Val Leu  
 20 25 30  
 Arg Gln Phe Gln Val Pro Lys Leu Gly Asn Gly Gly Gly Leu Gly Met  
 35 40 45  
 Val Ile Glu Cys Ser Ser Arg Pro Gln Lys Lys Ser Thr Ala His His  
 50 55 60  
 55 Arg Lys Thr Arg Pro Lys Lys Thr Gln Pro Trp Asp Ile Lys Arg Lys  
 65 70 75 80  
 Pro Thr Val Tyr Ala Pro Leu Pro Pro Leu Pro Ala Glu Trp Ser Pro  
 85 90 95  
 60 Phe Thr Leu Ala Ser Asn Asp Gly Gly Ala Ala Val Ala Ala Ser Pro  
 100 105 110  
 Ala Gly Asp Leu Val Ser Gly Ser Ala

534

115 120  
(2) INFORMATION FOR SEQ ID NO:406:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 109 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: peptide  
(B) LOCATION: 1..109  
(D) OTHER INFORMATION: / Ceres Seq. ID 2029081

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:406:

Met Ser Val Ser Ala Ile Phe Gly Thr Gly Ile Val Thr Val Ala Ala  
1 5 10 15  
Ser Pro Val Leu Arg Gln Phe Gln Val Pro Lys Leu Gly Asn Gly Gly  
20 25 30  
Gly Leu Gly Met Val Ile Glu Cys Ser Ser Arg Pro Gln Lys Lys Ser  
35 40 45  
Thr Ala His His Arg Lys Thr Arg Pro Lys Lys Thr Gln Pro Trp Asp  
50 55 60  
Ile Lys Arg Lys Pro Thr Val Tyr Ala Pro Leu Pro Pro Leu Pro Ala  
65 70 75 80  
Glu Trp Ser Pro Phe Thr Leu Ala Ser Asn Asp Gly Gly Ala Ala Val  
85 90 95  
Ala Ala Ser Pro Ala Gly Asp Leu Val Ser Gly Ser Ala  
100 105

(2) INFORMATION FOR SEQ ID NO:407:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 74 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: peptide  
(B) LOCATION: 1..74  
(D) OTHER INFORMATION: / Ceres Seq. ID 2029082

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:407:

Met Val Ile Glu Cys Ser Ser Arg Pro Gln Lys Lys Ser Thr Ala His  
1 5 10 15  
His Arg Lys Thr Arg Pro Lys Lys Thr Gln Pro Trp Asp Ile Lys Arg  
20 25 30  
Lys Pro Thr Val Tyr Ala Pro Leu Pro Pro Leu Pro Ala Glu Trp Ser  
35 40 45  
Pro Phe Thr Leu Ala Ser Asn Asp Gly Gly Ala Ala Val Ala Ala Ser  
50 55 60  
Pro Ala Gly Asp Leu Val Ser Gly Ser Ala  
65 70

(2) INFORMATION FOR SEQ ID NO:408:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 787 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: -  
(B) LOCATION: 1..787  
(D) OTHER INFORMATION: / Ceres Seq. ID 2029225

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:408:

agggattcta gggctttcct tcattgggtca gatcagacga cgtttttacat cttctttcttc

60

535

ttcctcttcg atattcgta gtgtgtgtat tttggggaaa actttgtgag caaagagcga 120  
 gaaaatgagc ggagccggta agaaaatcgc ggatgtggct ttcaaagctt caaggactat 180  
 cgattgggat ggtatggcta aggtccttgt cacagatgag gctcgtagag agttctctaa 240  
 ccttcgtcgt gctttcgatg aggttaacac acagctccag accaaattta gtcaggaacc 300  
 5 tgaacctata gattgggatt actataggaa gggatttgga gctggcattg ttgacaagta 360  
 caaggaagct tatgacagca ttgagattcc aaagtacgtt gacaaagtta ctcctgaata 420  
 caagccaaag tttgatgctt tggttggtgga actgaaagaa gcagaacaga aatcgtctca 480  
 ggagtctgaa cggtttgaga aagaaattgc tgatgtccaa gagatcagca aaaagctcag 540  
 caccatgact gcagatgagt actttgagaa gcacccggaa ctcaaaaaga agtttgatga 600  
 10 cgaaatccgt aatgacaact ggggatactg atcatgtttc tccatctccg gcttggaag 660  
 aaaactctct tctcttttct ctgttctctt actgtgattt tgtgagccaa tcataacaat 720  
 aataagtaca ccattcactt aagcagtgtt gagatcttca ttccaaggaa gataaacgca 780  
 tttgggtt

(2) INFORMATION FOR SEQ ID NO:409:

15 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 168 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..168

(D) OTHER INFORMATION: / Ceres Seq. ID 2029226

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:409:

Met Ser Gly Ala Gly Lys Lys Ile Ala Asp Val Ala Phe Lys Ala Ser  
 1 5 10 15  
 Arg Thr Ile Asp Trp Asp Gly Met Ala Lys Val Leu Val Thr Asp Glu  
 20 25 30  
 30 Ala Arg Arg Glu Phe Ser Asn Leu Arg Arg Ala Phe Asp Glu Val Asn  
 35 40 45  
 Thr Gln Leu Gln Thr Lys Phe Ser Gln Glu Pro Glu Pro Ile Asp Trp  
 50 55 60  
 35 Asp Tyr Tyr Arg Lys Gly Ile Gly Ala Gly Ile Val Asp Lys Tyr Lys  
 65 70 75 80  
 Glu Ala Tyr Asp Ser Ile Glu Ile Pro Lys Tyr Val Asp Lys Val Thr  
 85 90 95  
 Pro Glu Tyr Lys Pro Lys Phe Asp Ala Leu Leu Val Glu Leu Lys Glu  
 100 105 110  
 40 Ala Glu Gln Lys Ser Leu Lys Glu Ser Glu Arg Leu Glu Lys Glu Ile  
 115 120 125  
 Ala Asp Val Gln Glu Ile Ser Lys Lys Leu Ser Thr Met Thr Ala Asp  
 130 135 140  
 45 Glu Tyr Phe Glu Lys His Pro Glu Leu Lys Lys Lys Phe Asp Asp Glu  
 145 150 155 160  
 Ile Arg Asn Asp Asn Trp Gly Tyr  
 165

(2) INFORMATION FOR SEQ ID NO:410:

50 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 145 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

55 (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..145

(D) OTHER INFORMATION: / Ceres Seq. ID 2029227

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:410:

60 Met Ala Lys Val Leu Val Thr Asp Glu Ala Arg Arg Glu Phe Ser Asn  
 1 5 10 15  
 Leu Arg Arg Ala Phe Asp Glu Val Asn Thr Gln Leu Gln Thr Lys Phe

536

20 25 30  
 Ser Gln Glu Pro Glu Pro Ile Asp Trp Asp Tyr Tyr Arg Lys Gly Ile  
 35 40 45  
 Gly Ala Gly Ile Val Asp Lys Tyr Lys Glu Ala Tyr Asp Ser Ile Glu  
 5 50 55 60  
 Ile Pro Lys Tyr Val Asp Lys Val Thr Pro Glu Tyr Lys Pro Lys Phe  
 65 70 75 80  
 Asp Ala Leu Leu Val Glu Leu Lys Glu Ala Glu Gln Lys Ser Leu Lys  
 85 90 95  
 10 Glu Ser Glu Arg Leu Glu Lys Glu Ile Ala Asp Val Gln Glu Ile Ser  
 100 105 110  
 Lys Lys Leu Ser Thr Met Thr Ala Asp Glu Tyr Phe Glu Lys His Pro  
 115 120 125  
 Glu Leu Lys Lys Lys Phe Asp Asp Glu Ile Arg Asn Asp Asn Trp Gly  
 15 130 135 140  
 Tyr  
 145

(2) INFORMATION FOR SEQ ID NO:411:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1587 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..1587

(D) OTHER INFORMATION: / Ceres Seq. ID 2029422

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:411:

30 gtagtctctc tttttataac cacttctcga aaactgaaac ctttgtagag agaaccata 60  
 gttcgataaa acattctttt tgcaactgag acttggaac tgggttttac tcaaagtaag 120  
 atttcttctt gttgatggtg ataagttcaa tatgatgaaa agtttggtcta gtgcggttg 180  
 aggggaagacg gcgagggcat gtgatagctg cgtgaagagg cgggcacgtt ggtattgcgc 240  
 agctgatgat gccttttctt gccatgcttg tgacgggttc gtccactcgg caaacctct 300  
 35 tgctcgtagg cacgagagag ttcgcttgaa atcggctagc gccggaagt atcgccatgc 360  
 ctgcgccgct caccaagcca cgtggcatca gggatttaca cgtaaagctc ggacccacg 420  
 tggaggcaag aagagccaca cgatggtttt tcatgatctt gtgccggaga tgagcacgga 480  
 ggatcaagcg gagagttacg aggtggaaga gcagctcata tttagaggtgc cggatgatgaa 540  
 ctcgatgggt gaggagcaat gctttaacca atccctggag aaacagaatg agtttccaat 600  
 40 gatgccctta agtttcaaga gtagtgacga agaagatgac gacaacgctg agagttgtct 660  
 gaatggtttg tttccaaccg acatggaaact agctcagttc acagctgacg tggagactct 720  
 actcgggtgga ggggtcagag agtttcattc catagaagaa ctagggttag gtgagatgtt 780  
 aaagatcgaa aaagaggagg tggaggaaga gggagttgtg acaagagaag tgcgatgaca 840  
 agatgaaggt gatgagacat cccatttga aataagcttt gactacgagt acacacacaa 900  
 45 gaccacattc gatgaaggag aagaagatga gaaagaagac gtgatgaaga atgtgatgga 960  
 gatgggagtg aatgagatga gtgggtggat taaagaagag aagaaggaga aggtctcttat 1020  
 gcttagattg gactatgaat cagtcatttc cacttgggga ggccaaggga tcccatggac 1080  
 cgcccggtg ccatctgaaa tagacctcga catggtttgt tttccaacc ataccatggg 1140  
 tgaaagtgga gcagaggctc atcatcaca ccacttccgc ggcctagggt tacacctagg 1200  
 50 agatgctggg gatggaggaa gagaggctag ggtttcaaga taccgagaga aaaggaggac 1260  
 aaggttggtc tccaagaaga taaggtagca ggtacgtaaa ttgaatgcag agaaaaggcc 1320  
 tcgcatgaaa ggaaggttcg tcaagagatc ttcaattggt gttgctcact aaagaactta 1380  
 attaattatg gatattaaat tactttgctc tcatcttgct tttttgttgc tatagttttg 1440  
 gtgattgtta gctttctttt tctgcattca tagagaattt tgcacgtttt tgtgagctac 1500  
 55 gtatgtacat aaatatatca ccaaaaaatg tgactatctt gtaagcactg atttatatag 1560  
 tcgatataac gtgaattttg attgctg

(2) INFORMATION FOR SEQ ID NO:412:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 406 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

537

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..406

(D) OTHER INFORMATION: / Ceres Seq. ID 2029423

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:412:

5 Met Met Lys Ser Leu Ala Ser Ala Val Gly Gly Lys Thr Ala Arg Ala  
 1 5 10 15  
 10 Cys Asp Ser Cys Val Lys Arg Arg Ala Arg Trp Tyr Cys Ala Ala Asp  
 20 20 25 30  
 Asp Ala Phe Leu Cys His Ala Cys Asp Gly Ser Val His Ser Ala Asn  
 35 40 45  
 Pro Leu Ala Arg Arg His Glu Arg Val Arg Leu Lys Ser Ala Ser Ala  
 50 55 60  
 15 Gly Lys Tyr Arg His Ala Ser Pro Pro His Gln Ala Thr Trp His Gln  
 65 70 75 80  
 Gly Phe Thr Arg Lys Ala Arg Thr Pro Arg Gly Gly Lys Lys Ser His  
 85 90 95  
 20 Thr Met Val Phe His Asp Leu Val Pro Glu Met Ser Thr Glu Asp Gln  
 100 105 110  
 Ala Glu Ser Tyr Glu Val Glu Glu Gln Leu Ile Phe Glu Val Pro Val  
 115 120 125  
 Met Asn Ser Met Val Glu Glu Gln Cys Phe Asn Gln Ser Leu Glu Lys  
 130 135 140  
 25 Gln Asn Glu Phe Pro Met Met Pro Leu Ser Phe Lys Ser Ser Asp Glu  
 145 150 155 160  
 Glu Asp Asp Asp Asn Ala Glu Ser Cys Leu Asn Gly Leu Phe Pro Thr  
 165 170 175  
 30 Asp Met Glu Leu Ala Gln Phe Thr Ala Asp Val Glu Thr Leu Leu Gly  
 180 185 190  
 Gly Gly Asp Arg Glu Phe His Ser Ile Glu Glu Leu Gly Leu Gly Glu  
 195 200 205  
 Met Leu Lys Ile Glu Lys Glu Glu Val Glu Glu Glu Gly Val Val Thr  
 210 215 220  
 35 Arg Glu Val His Asp Gln Asp Glu Gly Asp Glu Thr Ser Pro Phe Glu  
 225 230 235 240  
 Ile Ser Phe Asp Tyr Glu Tyr Thr His Lys Thr Thr Phe Asp Glu Gly  
 245 250 255  
 40 Glu Glu Asp Glu Lys Glu Asp Val Met Lys Asn Val Met Glu Met Gly  
 260 265 270  
 Val Asn Glu Met Ser Gly Gly Ile Lys Glu Glu Lys Lys Glu Lys Ala  
 275 280 285  
 Leu Met Leu Arg Leu Asp Tyr Glu Ser Val Ile Ser Thr Trp Gly Gly  
 290 295 300  
 45 Gln Gly Ile Pro Trp Thr Ala Arg Val Pro Ser Glu Ile Asp Leu Asp  
 305 310 315 320  
 Met Val Cys Phe Pro Thr His Thr Met Gly Glu Ser Gly Ala Glu Ala  
 325 330 335  
 50 His His His Asn His Phe Arg Gly Leu Gly Leu His Leu Gly Asp Ala  
 340 345 350  
 Gly Asp Gly Gly Arg Glu Ala Arg Val Ser Arg Tyr Arg Glu Lys Arg  
 355 360 365  
 Arg Thr Arg Leu Phe Ser Lys Lys Ile Arg Tyr Glu Val Arg Lys Leu  
 370 375 380  
 55 Asn Ala Glu Lys Arg Pro Arg Met Lys Gly Arg Phe Val Lys Arg Ser  
 385 390 395 400  
 Ser Ile Gly Val Ala His  
 405

(2) INFORMATION FOR SEQ ID NO:413:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 405 amino acids

(B) TYPE: amino acid

538

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

5

(A) NAME/KEY: peptide

(B) LOCATION: 1..405

(D) OTHER INFORMATION: / Ceres Seq. ID 2029424

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:413:

10 Met Lys Ser Leu Ala Ser Ala Val Gly Gly Lys Thr Ala Arg Ala Cys  
 1 5 10 15  
 Asp Ser Cys Val Lys Arg Arg Ala Arg Trp Tyr Cys Ala Ala Asp Asp  
 20 25 30  
 Ala Phe Leu Cys His Ala Cys Asp Gly Ser Val His Ser Ala Asn Pro  
 35 40 45  
 15 Leu Ala Arg Arg His Glu Arg Val Arg Leu Lys Ser Ala Ser Ala Gly  
 50 55 60  
 Lys Tyr Arg His Ala Ser Pro Pro His Gln Ala Thr Trp His Gln Gly  
 65 70 75 80  
 Phe Thr Arg Lys Ala Arg Thr Pro Arg Gly Lys Lys Ser His Thr  
 20 85 90 95  
 Met Val Phe His Asp Leu Val Pro Glu Met Ser Thr Glu Asp Gln Ala  
 100 105 110  
 Glu Ser Tyr Glu Val Glu Glu Gln Leu Ile Phe Glu Val Pro Val Met  
 115 120 125  
 25 Asn Ser Met Val Glu Glu Gln Cys Phe Asn Gln Ser Leu Glu Lys Gln  
 130 135 140  
 Asn Glu Phe Pro Met Met Pro Leu Ser Phe Lys Ser Ser Asp Glu Glu  
 145 150 155 160  
 Asp Asp Asp Asn Ala Glu Ser Cys Leu Asn Gly Leu Phe Pro Thr Asp  
 30 165 170 175  
 Met Glu Leu Ala Gln Phe Thr Ala Asp Val Glu Thr Leu Leu Gly Gly  
 180 185 190  
 Gly Asp Arg Glu Phe His Ser Ile Glu Glu Leu Gly Leu Gly Glu Met  
 195 200 205  
 35 Leu Lys Ile Glu Lys Glu Glu Val Glu Glu Glu Gly Val Val Thr Arg  
 210 215 220  
 Glu Val His Asp Gln Asp Glu Gly Asp Glu Thr Ser Pro Phe Glu Ile  
 225 230 235 240  
 Ser Phe Asp Tyr Glu Tyr Thr His Lys Thr Thr Phe Asp Glu Gly Glu  
 40 245 250 255  
 Glu Asp Glu Lys Glu Asp Val Met Lys Asn Val Met Glu Met Gly Val  
 260 265 270  
 Asn Glu Met Ser Gly Gly Ile Lys Glu Glu Lys Lys Glu Lys Ala Leu  
 275 280 285  
 45 Met Leu Arg Leu Asp Tyr Glu Ser Val Ile Ser Thr Trp Gly Gly Gln  
 290 295 300  
 Gly Ile Pro Trp Thr Ala Arg Val Pro Ser Glu Ile Asp Leu Asp Met  
 305 310 315 320  
 Val Cys Phe Pro Thr His Thr Met Gly Glu Ser Gly Ala Glu Ala His  
 50 325 330 335  
 His His Asn His Phe Arg Gly Leu Gly Leu His Leu Gly Asp Ala Gly  
 340 345 350  
 Asp Gly Gly Arg Glu Ala Arg Val Ser Arg Tyr Arg Glu Lys Arg Arg  
 355 360 365  
 55 Thr Arg Leu Phe Ser Lys Lys Ile Arg Tyr Glu Val Arg Lys Leu Asn  
 370 375 380  
 Ala Glu Lys Arg Pro Arg Met Lys Gly Arg Phe Val Lys Arg Ser Ser  
 385 390 395 400  
 Ile Gly Val Ala His  
 60 405

(2) INFORMATION FOR SEQ ID NO:414:

(i) SEQUENCE CHARACTERISTICS:

539

(A) LENGTH: 309 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..309

(D) OTHER INFORMATION: / Ceres Seq. ID 2029425

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:414:

Met	Val	Phe	His	Asp	Leu	Val	Pro	Glu	Met	Ser	Thr	Glu	Asp	Gln	Ala
1				5				10						15	
Glu	Ser	Tyr	Glu	Val	Glu	Glu	Gln	Leu	Ile	Phe	Glu	Val	Pro	Val	Met
			20					25					30		
Asn	Ser	Met	Val	Glu	Glu	Gln	Cys	Phe	Asn	Gln	Ser	Leu	Glu	Lys	Gln
		35					40					45			
Asn	Glu	Phe	Pro	Met	Met	Pro	Leu	Ser	Phe	Lys	Ser	Ser	Asp	Glu	Glu
	50					55					60				
Asp	Asp	Asp	Asn	Ala	Glu	Ser	Cys	Leu	Asn	Gly	Leu	Phe	Pro	Thr	Asp
65					70					75				80	
Met	Glu	Leu	Ala	Gln	Phe	Thr	Ala	Asp	Val	Glu	Thr	Leu	Leu	Gly	Gly
				85					90					95	
Gly	Asp	Arg	Glu	Phe	His	Ser	Ile	Glu	Glu	Leu	Gly	Leu	Gly	Glu	Met
			100					105					110		
Leu	Lys	Ile	Glu	Lys	Glu	Glu	Val	Glu	Glu	Glu	Gly	Val	Val	Thr	Arg
		115					120					125			
Glu	Val	His	Asp	Gln	Asp	Glu	Gly	Asp	Glu	Thr	Ser	Pro	Phe	Glu	Ile
	130					135					140				
Ser	Phe	Asp	Tyr	Glu	Tyr	Thr	His	Lys	Thr	Thr	Phe	Asp	Glu	Gly	Glu
145					150					155				160	
Glu	Asp	Glu	Lys	Glu	Asp	Val	Met	Lys	Asn	Val	Met	Glu	Met	Gly	Val
			165					170						175	
Asn	Glu	Met	Ser	Gly	Gly	Ile	Lys	Glu	Glu	Lys	Lys	Glu	Lys	Ala	Leu
		180						185					190		
Met	Leu	Arg	Leu	Asp	Tyr	Glu	Ser	Val	Ile	Ser	Thr	Trp	Gly	Gly	Gln
		195					200					205			
Gly	Ile	Pro	Trp	Thr	Ala	Arg	Val	Pro	Ser	Glu	Ile	Asp	Leu	Asp	Met
	210					215					220				
Val	Cys	Phe	Pro	Thr	His	Thr	Met	Gly	Glu	Ser	Gly	Ala	Glu	Ala	His
225					230					235				240	
His	His	Asn	His	Phe	Arg	Gly	Leu	Gly	Leu	His	Leu	Gly	Asp	Ala	Gly
			245					250						255	
Asp	Gly	Gly	Arg	Glu	Ala	Arg	Val	Ser	Arg	Tyr	Arg	Glu	Lys	Arg	Arg
			260					265					270		
Thr	Arg	Leu	Phe	Ser	Lys	Lys	Ile	Arg	Tyr	Glu	Val	Arg	Lys	Leu	Asn
		275					280					285			
Ala	Glu	Lys	Arg	Pro	Arg	Met	Lys	Gly	Arg	Phe	Val	Lys	Arg	Ser	Ser
	290					295					300				

Ile Gly Val Ala His

50 305

(2) INFORMATION FOR SEQ ID NO:415:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1300 base pairs

(B) TYPE: nucleic acid

55 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

60 (B) LOCATION: 1..1300

(D) OTHER INFORMATION: / Ceres Seq. ID 2029806

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:415:

540

ctcaagtttc ttgcaattac ctacttacaa caaagcaact cgactcgaac aaacacatcc 60  
 atggctcctt ctgcgcaacc tcttcctgtg agtgtttcgg atgaaaaata tgcgaatgtc 120  
 aagtgggaag agttggcatt caagtttgtt cgtacggatt atatgtatgt tgcgaagtgc 180  
 aatcatggag agagttttca agaggggaag attcttcctt ttgctgattt gcaacttaac 240  
 5 ccttgcgctg ctgttcttca gtatggccag gggtttatatg aaggactgaa agcttacagg 300  
 acagaagatg gtcggattct gctattccga ccagacaaaa acggtctccg ccttcaagcc 360  
 ggagctgaca gactctatat gccttatcct tcggtcgcac aattcgtctc cgccatcaaa 420  
 caagttgctc ttgccaacaa gaaatggatt cctcctccgg ggaaaggaac attgtatat 480  
 aggcctatct tgtttgggag tggtcggatt cttggttcat ttccattcc tgagaccacc 540  
 10 ttcacagctt ttgcctgtcc tgttgacgt tatcataagg ataactctgg tttgaatctg 600  
 aaaatcgaag atcagtttctg tcgagctttt cctagtggaa ctggtggtgt gaagagcatc 660  
 acaaactatt gtcctgtttg gataccattg gcagaggcga aaaaacaagg tttctctgat 720  
 attttgtttt tggatgctgc aactggcaaa aacattgaag aacttttcgc agctaattgt 780  
 tttatgctca agggcaatgt tgtatcgaca ccaacaattg caggaactat tttgcccga 840  
 15 gtcactcgaa actgcgtaat ggaattgtgt cgtgatttcg gctaccagg caggaacgt 900  
 acgattcctc tagtggactt tctcgatgcg gacgaagctt tctgtactgg cactgcttcc 960  
 attgtgacta gtattgcac cgtaacctt aaagacaaaa agaccggatt caaaacaggg 1020  
 gaagaaacat tggctgcgaa gctataccag acgttaagt ataccagac gggtcgggtc 1080  
 gaggatacca agggatggac ggtggagatt gaccgccagg gctgaaagt gaaactgtaa 1140  
 20 cttgatgcta aatatgtgtg tgtgtgtata caaaacttat gtaagaaaca tctgaagtgt 1200  
 tctctgatct ttgtgattgt gatgatcatg catgtcctat cgttcgcatg tattttataa 1260  
 atgttcgtct gtaagttatt taataactat ggctttttgc

(2) INFORMATION FOR SEQ ID NO:416:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 374 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..374

(D) OTHER INFORMATION: / Ceres Seq. ID 2029807

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:416:

35 Leu Lys Phe Leu Ala Ile Thr Tyr Leu Gln Gln Ser Asn Ser Thr Arg  
 1 5 10 15  
 Thr Asn Thr Ser Met Ala Pro Ser Ala Gln Pro Leu Pro Val Ser Val  
 20 25 30  
 Ser Asp Glu Lys Tyr Ala Asn Val Lys Trp Glu Glu Leu Ala Phe Lys  
 40 35 40 45  
 Phe Val Arg Thr Asp Tyr Met Tyr Val Ala Lys Cys Asn His Gly Glu  
 50 55 60  
 Ser Phe Gln Glu Gly Lys Ile Leu Pro Phe Ala Asp Leu Gln Leu Asn  
 65 70 75 80  
 45 Pro Cys Ala Ala Val Leu Gln Tyr Gly Gln Gly Leu Tyr Glu Gly Leu  
 85 90 95  
 Lys Ala Tyr Arg Thr Glu Asp Gly Arg Ile Leu Leu Phe Arg Pro Asp  
 100 105 110  
 Gln Asn Gly Leu Arg Leu Gln Ala Gly Ala Asp Arg Leu Tyr Met Pro  
 50 115 120 125  
 Tyr Pro Ser Val Asp Gln Phe Val Ser Ala Ile Lys Gln Val Ala Leu  
 130 135 140  
 Ala Asn Lys Lys Trp Ile Pro Pro Pro Gly Lys Gly Thr Leu Tyr Ile  
 145 150 155 160  
 55 Arg Pro Ile Leu Phe Gly Ser Gly Pro Ile Leu Gly Ser Phe Pro Ile  
 165 170 175  
 Pro Glu Thr Thr Phe Thr Ala Phe Ala Cys Pro Val Gly Arg Tyr His  
 180 185 190  
 Lys Asp Asn Ser Gly Leu Asn Leu Lys Ile Glu Asp Gln Phe Arg Arg  
 60 195 200 205  
 Ala Phe Pro Ser Gly Thr Gly Gly Val Lys Ser Ile Thr Asn Tyr Cys  
 210 215 220



541

Pro Val Trp Ile Pro Leu Ala Glu Ala Lys Lys Gln Gly Phe Ser Asp  
 225 230 235 240  
 Ile Leu Phe Leu Asp Ala Ala Thr Gly Lys Asn Ile Glu Glu Leu Phe  
 245 250 255  
 5 Ala Ala Asn Val Phe Met Leu Lys Gly Asn Val Val Ser Thr Pro Thr  
 260 265 270  
 Ile Ala Gly Thr Ile Leu Pro Gly Val Thr Arg Asn Cys Val Met Glu  
 275 280 285  
 10 Leu Cys Arg Asp Phe Gly Tyr Gln Val Glu Glu Arg Thr Ile Pro Leu  
 290 295 300  
 Val Asp Phe Leu Asp Ala Asp Glu Ala Phe Cys Thr Gly Thr Ala Ser  
 305 310 315 320  
 Ile Val Thr Ser Ile Ala Ser Val Thr Phe Lys Asp Lys Lys Thr Gly  
 325 330 335  
 15 Phe Lys Thr Gly Glu Glu Thr Leu Ala Ala Lys Leu Tyr Glu Thr Leu  
 340 345 350  
 Ser Asp Ile Gln Thr Gly Arg Val Glu Asp Thr Lys Gly Trp Thr Val  
 355 360 365  
 20 Glu Ile Asp Arg Gln Gly  
 370  
 (2) INFORMATION FOR SEQ ID NO:417:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 354 amino acids  
 (B) TYPE: amino acid  
 25 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 30 (B) LOCATION: 1..354  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2029808  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:417:  
 Met Ala Pro Ser Ala Gln Pro Leu Pro Val Ser Val Ser Asp Glu Lys  
 1 5 10 15  
 35 Tyr Ala Asn Val Lys Trp Glu Glu Leu Ala Phe Lys Phe Val Arg Thr  
 20 25 30  
 Asp Tyr Met Tyr Val Ala Lys Cys Asn His Gly Glu Ser Phe Gln Glu  
 35 40 45  
 40 Gly Lys Ile Leu Pro Phe Ala Asp Leu Gln Leu Asn Pro Cys Ala Ala  
 50 55 60  
 Val Leu Gln Tyr Gly Gln Gly Leu Tyr Glu Gly Leu Lys Ala Tyr Arg  
 65 70 75 80  
 Thr Glu Asp Gly Arg Ile Leu Leu Phe Arg Pro Asp Gln Asn Gly Leu  
 85 90 95  
 45 Arg Leu Gln Ala Gly Ala Asp Arg Leu Tyr Met Pro Tyr Pro Ser Val  
 100 105 110  
 Asp Gln Phe Val Ser Ala Ile Lys Gln Val Ala Leu Ala Asn Lys Lys  
 115 120 125  
 50 Trp Ile Pro Pro Pro Gly Lys Gly Thr Leu Tyr Ile Arg Pro Ile Leu  
 130 135 140  
 Phe Gly Ser Gly Pro Ile Leu Gly Ser Phe Pro Ile Pro Glu Thr Thr  
 145 150 155 160  
 Phe Thr Ala Phe Ala Cys Pro Val Gly Arg Tyr His Lys Asp Asn Ser  
 165 170 175  
 55 Gly Leu Asn Leu Lys Ile Glu Asp Gln Phe Arg Arg Ala Phe Pro Ser  
 180 185 190  
 Gly Thr Gly Gly Val Lys Ser Ile Thr Asn Tyr Cys Pro Val Trp Ile  
 195 200 205  
 60 Pro Leu Ala Glu Ala Lys Lys Gln Gly Phe Ser Asp Ile Leu Phe Leu  
 210 215 220  
 Asp Ala Ala Thr Gly Lys Asn Ile Glu Glu Leu Phe Ala Ala Asn Val  
 225 230 235 240

542

	Phe	Met	Leu	Lys	Gly	Asn	Val	Val	Ser	Thr	Pro	Thr	Ile	Ala	Gly	Thr
					245					250					255	
	Ile	Leu	Pro	Gly	Val	Thr	Arg	Asn	Cys	Val	Met	Glu	Leu	Cys	Arg	Asp
				260					265					270		
5	Phe	Gly	Tyr	Gln	Val	Glu	Glu	Arg	Thr	Ile	Pro	Leu	Val	Asp	Phe	Leu
			275					280					285			
	Asp	Ala	Asp	Glu	Ala	Phe	Cys	Thr	Gly	Thr	Ala	Ser	Ile	Val	Thr	Ser
		290					295					300				
	Ile	Ala	Ser	Val	Thr	Phe	Lys	Asp	Lys	Lys	Thr	Gly	Phe	Lys	Thr	Gly
10	305					310					315					320
	Glu	Glu	Thr	Leu	Ala	Ala	Lys	Leu	Tyr	Glu	Thr	Leu	Ser	Asp	Ile	Gln
				325						330					335	
	Thr	Gly	Arg	Val	Glu	Asp	Thr	Lys	Gly	Trp	Thr	Val	Glu	Ile	Asp	Arg
				340					345						350	
15	Gln	Gly														

## (2) INFORMATION FOR SEQ ID NO:418:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 320 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

(A) NAME/KEY: peptide  
 (B) LOCATION: 1..320

(D) OTHER INFORMATION: / Ceres Seq. ID 2029809

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:418:

30	Met	Tyr	Val	Ala	Lys	Cys	Asn	His	Gly	Glu	Ser	Phe	Gln	Glu	Gly	Lys
	1			5					10					15		
	Ile	Leu	Pro	Phe	Ala	Asp	Leu	Gln	Leu	Asn	Pro	Cys	Ala	Ala	Val	Leu
				20					25				30			
	Gln	Tyr	Gly	Gln	Gly	Leu	Tyr	Glu	Gly	Leu	Lys	Ala	Tyr	Arg	Thr	Glu
			35				40					45				
35	Asp	Gly	Arg	Ile	Leu	Leu	Phe	Arg	Pro	Asp	Gln	Asn	Gly	Leu	Arg	Leu
		50					55				60					
	Gln	Ala	Gly	Ala	Asp	Arg	Leu	Tyr	Met	Pro	Tyr	Pro	Ser	Val	Asp	Gln
	65				70					75					80	
	Phe	Val	Ser	Ala	Ile	Lys	Gln	Val	Ala	Leu	Ala	Asn	Lys	Lys	Trp	Ile
40					85					90					95	
	Pro	Pro	Pro	Gly	Lys	Gly	Thr	Leu	Tyr	Ile	Arg	Pro	Ile	Leu	Phe	Gly
				100					105					110		
	Ser	Gly	Pro	Ile	Leu	Gly	Ser	Phe	Pro	Ile	Pro	Glu	Thr	Thr	Phe	Thr
			115					120					125			
45	Ala	Phe	Ala	Cys	Pro	Val	Gly	Arg	Tyr	His	Lys	Asp	Asn	Ser	Gly	Leu
		130					135					140				
	Asn	Leu	Lys	Ile	Glu	Asp	Gln	Phe	Arg	Arg	Ala	Phe	Pro	Ser	Gly	Thr
	145					150					155					160
	Gly	Gly	Val	Lys	Ser	Ile	Thr	Asn	Tyr	Cys	Pro	Val	Trp	Ile	Pro	Leu
50					165					170					175	
	Ala	Glu	Ala	Lys	Lys	Gln	Gly	Phe	Ser	Asp	Ile	Leu	Phe	Leu	Asp	Ala
				180					185					190		
	Ala	Thr	Gly	Lys	Asn	Ile	Glu	Glu	Leu	Phe	Ala	Ala	Asn	Val	Phe	Met
			195					200					205			
55	Leu	Lys	Gly	Asn	Val	Val	Ser	Thr	Pro	Thr	Ile	Ala	Gly	Thr	Ile	Leu
		210					215					220				
	Pro	Gly	Val	Thr	Arg	Asn	Cys	Val	Met	Glu	Leu	Cys	Arg	Asp	Phe	Gly
	225					230					235					240
	Tyr	Gln	Val	Glu	Glu	Arg	Thr	Ile	Pro	Leu	Val	Asp	Phe	Leu	Asp	Ala
				245						250					255	
60	Asp	Glu	Ala	Phe	Cys	Thr	Gly	Thr	Ala	Ser	Ile	Val	Thr	Ser	Ile	Ala
				260					265						270	

543

Ser Val Thr Phe Lys Asp Lys Lys Thr Gly Phe Lys Thr Gly Glu Glu  
 275 280 285  
 Thr Leu Ala Ala Lys Leu Tyr Glu Thr Leu Ser Asp Ile Gln Thr Gly  
 290 295 300  
 Arg Val Glu Asp Thr Lys Gly Trp Thr Val Glu Ile Asp Arg Gln Gly  
 305 310 315 320

## (2) INFORMATION FOR SEQ ID NO:419:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 987 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..987

(D) OTHER INFORMATION: / Ceres Seq. ID 2030038

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:419:

```

atatctgaaa gactcaaaaa cctatcgtca tttatcttca acaggcggaa taacggagat      60
ggccgccatt acagctctca ctctccgctc tcctgtttat ctcccatcat ctgccacaag      120
ccctagattc catggcttca ccaatcaacc accaccagct cgtctcttct ttcctcttaa      180
ccccctccct tctctatcaa tccaaaaccc taaatccatc cgaatttccg catccgcttc      240
gcccataaca acaccaatcc tccagacgga gaagtcaaca gtcgggtcat cgacactcac      300
cggtgccact cggtctctcg ctactctcgc ggcttttagca atcgctgtaa ccagagtcct      360
agctcagaaa ctctctctcg caatccagac ttcaagtcct gtaatcgccg acgggttacg      420
attctctctc agtacgcgcg gacctgtctt ctccgctctt ctccgggata gtcctccggg      480
atacttgaac acgccgctga cggttgttgc ggtggggata aagaagtggc tagacattta      540
cagtggggta ttgatggta gggttttgct gagttgggtc cctaatatcc cttgggaaag      600
acagcctttg tctgccatta gagatctctg tgatccttac ttgaatctct tcagaaacat      660
cattcctcct atcttcgata cgcttgatgt tagtccattg cctgctttcg cggttctttg      720
tacacttgga tcgattgttc atggcagcac ttggtagaaa ttggaagact tgagctaata      780
gtttttgttg aacgttcaat tgggaaattt ctgcaacttg tttccctaga ttagggaaaa      840
atcagaattt gatgctgtgt tggatgctcg atgagagttg tgagctt at cgtttttgtt      900
gttctctttt tgactttgat tctgaataat gagattttgg gttgttgtt atctcctaata      960
tacatgttca gactaccgtt tattgat

```

## (2) INFORMATION FOR SEQ ID NO:420:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 251 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..251

(D) OTHER INFORMATION: / Ceres Seq. ID 2030039

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:420:

```

Tyr Leu Lys Asp Ser Lys Thr Tyr Arg His Leu Ser Ser Thr Gly Gly
1      5      10      15
Ile Thr Glu Met Ala Ala Ile Thr Ala Leu Thr Leu Arg Ser Pro Val
20      25      30
Tyr Leu Pro Ser Ser Ala Thr Ser Pro Arg Phe His Gly Phe Thr Asn
35      40      45
Gln Pro Pro Pro Ala Arg Leu Phe Phe Pro Leu Asn Pro Phe Pro Ser
50      55      60
Leu Ser Ile Gln Asn Pro Lys Ser Ile Arg Ile Ser Ala Ser Ala Ser
65      70      75      80
Pro Ile Thr Thr Pro Ile Leu Gln Thr Glu Lys Ser Thr Ala Arg Ser
85      90      95
Ser Thr Leu Thr Gly Ser Thr Arg Ser Leu Ala Thr Leu Ala Ala Leu
100      105      110
Ala Ile Ala Val Thr Arg Val Leu Ala Gln Lys Leu Ser Leu Ala Ile

```

544

115 120 125  
 Gln Thr Ser Ser Pro Val Ile Ala Asp Gly Leu Arg Phe Ser Leu Ser  
 130 135 140  
 Thr Ala Gly Pro Val Phe Phe Ala Ser Leu Arg Asp Arg Pro Pro Gly  
 145 150 155 160  
 Tyr Leu Asn Thr Pro Leu Thr Val Val Ala Val Gly Ile Lys Lys Trp  
 165 170 175  
 Leu Asp Ile Tyr Ser Gly Val Leu Met Val Arg Val Leu Leu Ser Trp  
 180 185 190  
 10 Phe Pro Asn Ile Pro Trp Glu Arg Gln Pro Leu Ser Ala Ile Arg Asp  
 195 200 205  
 Leu Cys Asp Pro Tyr Leu Asn Leu Phe Arg Asn Ile Ile Pro Pro Ile  
 210 215 220  
 Phe Asp Thr Leu Asp Val Ser Pro Leu Leu Ala Phe Ala Val Leu Gly  
 15 225 230 235 240  
 Thr Leu Gly Ser Ile Val His Gly Ser Thr Gly  
 245 250

(2) INFORMATION FOR SEQ ID NO:421:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 232 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

25 (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..232

(D) OTHER INFORMATION: / Ceres Seq. ID 2030040

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:421:

30 Met Ala Ala Ile Thr Ala Leu Thr Leu Arg Ser Pro Val Tyr Leu Pro  
 1 5 10 15  
 Ser Ser Ala Thr Ser Pro Arg Phe His Gly Phe Thr Asn Gln Pro Pro  
 20 25 30  
 Pro Ala Arg Leu Phe Phe Pro Leu Asn Pro Phe Pro Ser Leu Ser Ile  
 35 35 40 45  
 Gln Asn Pro Lys Ser Ile Arg Ile Ser Ala Ser Ala Ser Pro Ile Thr  
 50 55 60  
 Thr Pro Ile Leu Gln Thr Glu Lys Ser Thr Ala Arg Ser Ser Thr Leu  
 65 70 75 80  
 40 Thr Gly Ser Thr Arg Ser Leu Ala Thr Leu Ala Ala Leu Ala Ile Ala  
 85 90 95  
 Val Thr Arg Val Leu Ala Gln Lys Leu Ser Leu Ala Ile Gln Thr Ser  
 100 105 110  
 Ser Pro Val Ile Ala Asp Gly Leu Arg Phe Ser Leu Ser Thr Ala Gly  
 115 120 125  
 45 Pro Val Phe Phe Ala Ser Leu Arg Asp Arg Pro Pro Gly Tyr Leu Asn  
 130 135 140  
 Thr Pro Leu Thr Val Val Ala Val Gly Ile Lys Lys Trp Leu Asp Ile  
 145 150 155 160  
 50 Tyr Ser Gly Val Leu Met Val Arg Val Leu Leu Ser Trp Phe Pro Asn  
 165 170 175  
 Ile Pro Trp Glu Arg Gln Pro Leu Ser Ala Ile Arg Asp Leu Cys Asp  
 180 185 190  
 Pro Tyr Leu Asn Leu Phe Arg Asn Ile Ile Pro Pro Ile Phe Asp Thr  
 195 200 205  
 55 Leu Asp Val Ser Pro Leu Leu Ala Phe Ala Val Leu Gly Thr Leu Gly  
 210 215 220  
 Ser Ile Val His Gly Ser Thr Gly  
 225 230

60 (2) INFORMATION FOR SEQ ID NO:422:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 755 base pairs

545

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

5

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..755

(D) OTHER INFORMATION: / Ceres Seq. ID 2031778

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:422:

```

10  ccgcagaaaa ccctattgct cccctttgct caagctcagt gcctctttgc agcgaaagct      60
    caaccaacaa gacgatgatt atctcagaga acaatcgag agagatctgc aaataccttt      120
    tcaaagaggg agtatgcttt gctaagaagg atttcaatct cccaaagcat ccggttgattg      180
    atgtaccaa cctgcaagtg attaagctca tgcagagttt caagtccaag gagtacgtta      240
    gggagacatt tgcgtggatg cattattatt ggtttctgac taatgaagga attgagttct      300
15  tgagaactta tcttaacctt ccttccgatg ttgtccctgc tactttgaag aagtctgcta      360
    agcccgtggt tgcgtccctt ggtggccac ctggtgatcg ccaaagagga ccacctcgct      420
    ctgatggaga ccgtcccaga tttggtgacc gtgatggata ccgtggaggc ccacgtggtg      480
    gtgatgagaa ggggtggagct ccagctgatt tccagccgtc tttccaagga ggtggtggtg      540
    ggcctggttt tggccgtggt gcaggcgggt acagtgcagc agcaccatct ggttcagggt      600
20  tcccttgaaa aatttgttgt catattgcga cgatggaagg acagttttgt tttttgttct      660
    agttctgttt gtgtaatgca aatctggaat ctataatcta totattacct tcaccttggt      720
    ttaaacgaac aaatcctgta gtttcaggat tagtg

```

(2) INFORMATION FOR SEQ ID NO:423:

(i) SEQUENCE CHARACTERISTICS:

25

(A) LENGTH: 201 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

30

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..201

(D) OTHER INFORMATION: / Ceres Seq. ID 2031779

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:423:

```

35  Ala Glu Asn Pro Ile Ala Pro Leu Cys Ser Ser Ser Val Pro Leu Cys
    1          5          10          15
    Ser Glu Ser Ser Thr Asn Lys Thr Met Ile Ile Ser Glu Asn Asn Arg
        20          25          30
    Arg Glu Ile Cys Lys Tyr Leu Phe Lys Glu Gly Val Cys Phe Ala Lys
    40          35          40          45
    Lys Asp Phe Asn Leu Pro Lys His Pro Leu Ile Asp Val Pro Asn Leu
        50          55          60
    Gln Val Ile Lys Leu Met Gln Ser Phe Lys Ser Lys Glu Tyr Val Arg
        65          70          75          80
45  Glu Thr Phe Ala Trp Met His Tyr Tyr Trp Phe Leu Thr Asn Glu Gly
        85          90          95
    Ile Glu Phe Leu Arg Thr Tyr Leu Asn Leu Pro Ser Asp Val Val Pro
        100          105          110
    Ala Thr Leu Lys Lys Ser Ala Lys Pro Gly Gly Arg Pro Phe Gly Gly
    50          115          120          125
    Pro Pro Gly Asp Arg Gln Arg Gly Pro Pro Arg Ser Asp Gly Asp Arg
        130          135          140
    Pro Arg Phe Gly Asp Arg Asp Gly Tyr Arg Gly Gly Pro Arg Gly Gly
        145          150          155          160
55  Asp Glu Lys Gly Gly Ala Pro Ala Asp Phe Gln Pro Ser Phe Gln Gly
        165          170          175
    Gly Gly Gly Arg Pro Gly Phe Gly Arg Gly Ala Gly Gly Tyr Ser Ala
        180          185          190
    Ala Ala Pro Ser Gly Ser Gly Phe Pro
        195          200

```

(2) INFORMATION FOR SEQ ID NO:424:

(i) SEQUENCE CHARACTERISTICS:

546

(A) LENGTH: 177 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..177

(D) OTHER INFORMATION: / Ceres Seq. ID 2031780

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:424:

```

Met Ile Ile Ser Glu Asn Asn Arg Arg Glu Ile Cys Lys Tyr Leu Phe
 1             5             10             15
Lys Glu Gly Val Cys Phe Ala Lys Lys Asp Phe Asn Leu Pro Lys His
          20             25             30
15 Pro Leu Ile Asp Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln Ser
          35             40             45
Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His Tyr
          50             55             60
20 Tyr Trp Phe Leu Thr Asn Glu Gly Ile Glu Phe Leu Arg Thr Tyr Leu
          65             70             75             80
Asn Leu Pro Ser Asp Val Val Pro Ala Thr Leu Lys Lys Ser Ala Lys
          85             90             95
Pro Gly Gly Arg Pro Phe Gly Gly Pro Pro Gly Asp Arg Gln Arg Gly
          100            105            110
25 Pro Pro Arg Ser Asp Gly Asp Arg Pro Arg Phe Gly Asp Arg Asp Gly
          115            120            125
Tyr Arg Gly Gly Pro Arg Gly Gly Asp Glu Lys Gly Gly Ala Pro Ala
          130            135            140
30 Asp Phe Gln Pro Ser Phe Gln Gly Gly Gly Gly Arg Pro Gly Phe Gly
          145            150            155            160
Arg Gly Ala Gly Gly Tyr Ser Ala Ala Ala Pro Ser Gly Ser Gly Phe
          165            170            175
Pro

```

35 (2) INFORMATION FOR SEQ ID NO:425:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 132 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

40 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..132

(D) OTHER INFORMATION: / Ceres Seq. ID 2031781

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:425:

```

Met Gln Ser Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp
 1             5             10             15
Met His Tyr Tyr Trp Phe Leu Thr Asn Glu Gly Ile Glu Phe Leu Arg
          20             25             30
50 Thr Tyr Leu Asn Leu Pro Ser Asp Val Val Pro Ala Thr Leu Lys Lys
          35             40             45
Ser Ala Lys Pro Gly Gly Arg Pro Phe Gly Gly Pro Pro Gly Asp Arg
          50             55             60
55 Gln Arg Gly Pro Pro Arg Ser Asp Gly Asp Arg Pro Arg Phe Gly Asp
          65             70             75             80
Arg Asp Gly Tyr Arg Gly Gly Pro Arg Gly Gly Asp Glu Lys Gly Gly
          85             90             95
Ala Pro Ala Asp Phe Gln Pro Ser Phe Gln Gly Gly Gly Gly Arg Pro
          100            105            110
60 Gly Phe Gly Arg Gly Ala Gly Gly Tyr Ser Ala Ala Ala Pro Ser Gly
          115            120            125

```

547

Ser Gly Phe Pro  
130

## (2) INFORMATION FOR SEQ ID NO:426:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 538 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

- (A) NAME/KEY: -  
(B) LOCATION: 1..538  
(D) OTHER INFORMATION: / Ceres Seq. ID 2032723

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:426:

```

15 aaagcaataa aaaacaaaaa tggcggtttac tccgaagatc atcacatgcc tcattgtcct      60
   tacgatctac atggcatccc caacagagtc aaccatccag tgtgggacag tgacgagcac      120
   actggcacag tgcctgacct acttgaccaa cagtgggtcca ttgccatcac aatgctgcgt      180
   gggagtcaag tcattgtacc aattggctca gaccacaccg gaccgtaaac aagtatgtga      240
   gtgccttaaa ctagcgggta aagaaatcaa gggcctcaac accgaccttg tggccgcact      300
20 tcctaccact tgtggtgttt caattcccta cccatcagt tttagcacca attgcgacag      360
   tatatcgact gccgtgtgaa agaggctagt gatcagatgt acgactaatc aaacttgcca      420
   gcttttaacc taattaaata aaagtattct gcttatattt cccattttat gattttatct      480
   tcttatctat gtaaccacac gatttcatat gctaataatg acaacggatc tttctctc

```

## (2) INFORMATION FOR SEQ ID NO:427:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 119 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

- (A) NAME/KEY: peptide  
(B) LOCATION: 1..119  
(D) OTHER INFORMATION: / Ceres Seq. ID 2032724

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:427:

```

35 Met Ala Phe Thr Pro Lys Ile Ile Thr Cys Leu Ile Val Leu Thr Ile
   1           5           10           15
   Tyr Met Ala Ser Pro Thr Glu Ser Thr Ile Gln Cys Gly Thr Val Thr
           20           25           30
40 Ser Thr Leu Ala Gln Cys Leu Thr Tyr Leu Thr Asn Ser Gly Pro Leu
           35           40           45
   Pro Ser Gln Cys Cys Val Gly Val Lys Ser Leu Tyr Gln Leu Ala Gln
           50           55           60
   Thr Thr Pro Asp Arg Lys Gln Val Cys Glu Cys Leu Lys Leu Ala Gly
45 65           70           75           80
   Lys Glu Ile Lys Gly Leu Asn Thr Asp Leu Val Ala Ala Leu Pro Thr
           85           90           95
   Thr Cys Gly Val Ser Ile Pro Tyr Pro Ile Ser Phe Ser Thr Asn Cys
           100          105          110
50 Asp Ser Ile Ser Thr Ala Val
   115

```

## (2) INFORMATION FOR SEQ ID NO:428:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 102 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

- (A) NAME/KEY: peptide  
(B) LOCATION: 1..102  
(D) OTHER INFORMATION: / Ceres Seq. ID 2032725

548

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:428:

Met Ala Ser Pro Thr Glu Ser Thr Ile Gln Cys Gly Thr Val Thr Ser  
 1 5 10 15  
 Thr Leu Ala Gln Cys Leu Thr Tyr Leu Thr Asn Ser Gly Pro Leu Pro  
 20 25 30  
 Ser Gln Cys Cys Val Gly Val Lys Ser Leu Tyr Gln Leu Ala Gln Thr  
 35 40 45  
 Thr Pro Asp Arg Lys Gln Val Cys Glu Cys Leu Lys Leu Ala Gly Lys  
 50 55 60  
 Glu Ile Lys Gly Leu Asn Thr Asp Leu Val Ala Ala Leu Pro Thr Thr  
 65 70 75 80  
 Cys Gly Val Ser Ile Pro Tyr Pro Ile Ser Phe Ser Thr Asn Cys Asp  
 85 90 95  
 Ser Ile Ser Thr Ala Val  
 100

(2) INFORMATION FOR SEQ ID NO:429:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 75 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..75

(D) OTHER INFORMATION: / Ceres Seq. ID 2032726

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:429:

Lys Ala Ile Lys Asn Lys Asn Gly Val Tyr Ser Glu Asp His His Met  
 1 5 10 15  
 Pro His Cys Pro Tyr Asp Leu His Gly Ile Pro Asn Arg Val Asn His  
 20 25 30  
 Pro Val Trp Asp Ser Asp Glu His Thr Gly Thr Val Pro Asp Leu Leu  
 35 40 45  
 Asp Gln Gln Trp Ser Ile Ala Ile Thr Met Leu Arg Gly Ser Gln Val  
 50 55 60  
 Ile Val Pro Ile Gly Ser Asp His Thr Gly Pro  
 65 70 75

(2) INFORMATION FOR SEQ ID NO:430:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 198 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..198

(D) OTHER INFORMATION: / Ceres Seq. ID 2035536

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:430:

atgtcgttgg tatggttggga agcgatgttg cctctcggaa tcatcggtgg gatgctctgt 60  
 atcatgggca attctcagta ctacatccac aaagcttatc atggccgtcc taagcacatc 120  
 ggccacgatg aatgggatgt tgctatggaa agacgcgaca agaaagtcgt cgagaaagct 180  
 gcagctcctt cctcatga

(2) INFORMATION FOR SEQ ID NO:431:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 65 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide



549

(B) LOCATION: 1..65

(D) OTHER INFORMATION: / Ceres Seq. ID 2035537

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:431:

5 Met Ser Leu Val Trp Leu Glu Ala Met Leu Pro Leu Gly Ile Ile Gly  
 1 5 10 15  
 Gly Met Leu Cys Ile Met Gly Asn Ser Gln Tyr Tyr Ile His Lys Ala  
 20 25 30  
 Tyr His Gly Arg Pro Lys His Ile Gly His Asp Glu Trp Asp Val Ala  
 35 40 45  
 10 Met Glu Arg Arg Asp Lys Lys Val Val Glu Lys Ala Ala Ala Pro Ser  
 50 55 60  
 Ser  
 65

(2) INFORMATION FOR SEQ ID NO:432:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 65 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..65

(D) OTHER INFORMATION: / Ceres Seq. ID 2035538

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:432:

25 Cys Arg Trp Tyr Gly Trp Lys Arg Cys Cys Leu Ser Glu Ser Ser Val  
 1 5 10 15  
 Gly Cys Ser Val Ser Trp Ala Ile Leu Ser Thr Thr Ser Thr Lys Leu  
 20 25 30  
 30 Ile Met Ala Val Leu Ser Thr Ser Ala Thr Met Asn Gly Met Leu Leu  
 35 40 45  
 Trp Lys Asp Ala Thr Arg Lys Ser Ser Arg Lys Leu Gln Leu Leu Pro  
 50 55 60  
 His  
 65

(2) INFORMATION FOR SEQ ID NO:433:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 57 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..57

(D) OTHER INFORMATION: / Ceres Seq. ID 2035539

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:433:

50 Met Leu Pro Leu Gly Ile Ile Gly Gly Met Leu Cys Ile Met Gly Asn  
 1 5 10 15  
 Ser Gln Tyr Tyr Ile His Lys Ala Tyr His Gly Arg Pro Lys His Ile  
 20 25 30  
 Gly His Asp Glu Trp Asp Val Ala Met Glu Arg Arg Asp Lys Lys Val  
 35 40 45  
 55 Val Glu Lys Ala Ala Ala Pro Ser Ser  
 50 55

(2) INFORMATION FOR SEQ ID NO:434:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 573 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

550

## (ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..573

(D) OTHER INFORMATION: / Ceres Seq. ID 2035575

## 5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:434:

atggagactt	ctatgaggt	taccagcaat	tccaagtcta	tgaagattca	tgccaaagag	60
aagggtccgg	tgaactcaaa	aacccattta	cagcttcacg	gagagttaga	tactggaact	120
ggggctccga	gttacttctg	tgcatgatt	agacactttt	ttcctgaggc	ttcaacaggc	180
cttggggtag	gattgcatta	tgataagcgc	caaaagcttc	ggtgtcttgt	acgcggaaaa	240
aaagagtttc	ctgtaagagc	tgataagcgt	gtaaccttta	atattaaagg	gcggtgtgat	300
attgatcagg	acttaaata	gaagaacccc	aaaggagcag	cagaatttgc	ctggaacata	360
atggatttca	aggaagatca	ggatgtacgg	atcaaagttg	gctacgaaat	gtttgataag	420
gtcccttata	tgacagattag	agaaaacaat	tggaactctca	acgcgaacat	gaagggaaaa	480
tggaacttgc	ggtatgacct	gtaactgcat	ttttttcaat	catcatctga	gaaatgtatt	540
gataccactg	ctgatgaaca	catttttaatt	cta			

## 15 (2) INFORMATION FOR SEQ ID NO:435:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 167 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..167

(D) OTHER INFORMATION: / Ceres Seq. ID 2035576

## 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:435:

Met	Glu	Thr	Ser	Met	Arg	Tyr	Thr	Ser	Asn	Ser	Lys	Ser	Met	Lys	Ile	
1				5					10					15		
His	Ala	Lys	Glu	Lys	Val	Pro	Val	Asn	Ser	Lys	Thr	His	Leu	Gln	Leu	
				20					25					30		
His	Gly	Glu	Leu	Asp	Thr	Gly	Thr	Gly	Ala	Pro	Ser	Tyr	Phe	Cys	Ala	
				35					40					45		
Met	Ile	Arg	His	Phe	Phe	Pro	Glu	Ala	Ser	Thr	Gly	Leu	Gly	Val	Gly	
Leu	His	Tyr	Asp	Lys	Arg	Gln	Lys	Leu	Arg	Cys	Leu	Val	Arg	Gly	Lys	
Lys	Glu	Phe	Pro	Val	Arg	Ala	Asp	Lys	Arg	Val	Thr	Phe	Asn	Ile	Lys	
Gly	Arg	Cys	Asp	Ile	Asp	Gln	Asp	Leu	Asn	Gln	Lys	Asn	Pro	Lys	Gly	
Ala	Ala	Glu	Phe	Ala	Trp	Asn	Ile	Met	Asp	Phe	Lys	Glu	Asp	Gln	Asp	
Val	Arg	Ile	Lys	Val	Gly	Tyr	Glu	Met	Phe	Asp	Lys	Val	Pro	Tyr	Met	
Gln	Ile	Arg	Glu	Asn	Asn	Trp	Thr	Leu	Asn	Ala	Asn	Met	Lys	Gly	Lys	
Trp	Asn	Leu	Arg	Tyr	Asp	Leu										

## 50 (2) INFORMATION FOR SEQ ID NO:436:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 163 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..163

(D) OTHER INFORMATION: / Ceres Seq. ID 2035577

## 60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:436:

Met	Arg	Tyr	Thr	Ser	Asn	Ser	Lys	Ser	Met	Lys	Ile	His	Ala	Lys	Glu	
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	--

551

	1		5		10		15
	Lys	Val	Pro	Val	Asn	Ser	Lys
			20			25	
5	Asp	Thr	Gly	Thr	Gly	Ala	Pro
			35			40	
	Phe	Phe	Pro	Glu	Ala	Ser	Thr
			50			55	
	Lys	Arg	Gln	Lys	Leu	Arg	Cys
			65			70	
10	Val	Arg	Ala	Asp	Lys	Arg	Val
				85		90	
	Ile	Asp	Gln	Asp	Leu	Asn	Gln
			100			105	
	Ala	Trp	Asn	Ile	Met	Asp	Phe
15			115			120	
	Val	Gly	Tyr	Glu	Met	Phe	Asp
			130			135	
	Asn	Asn	Trp	Thr	Leu	Asn	Ala
			145			150	
20	Tyr	Asp	Leu				

## (2) INFORMATION FOR SEQ ID NO:437:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 154 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..154

(D) OTHER INFORMATION: / Ceres Seq. ID 2035578

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:437:

	Met	Lys	Ile	His	Ala	Lys	Glu	Lys	Val	Pro	Val	Asn	Ser	Lys	Thr	His
35	1				5				10					15		
	Leu	Gln	Leu	His	Gly	Glu	Leu	Asp	Thr	Gly	Thr	Gly	Ala	Pro	Ser	Tyr
			20					25					30			
	Phe	Cys	Ala	Met	Ile	Arg	His	Phe	Pro	Glu	Ala	Ser	Thr	Gly	Leu	
			35				40					45				
40	Gly	Val	Gly	Leu	His	Tyr	Asp	Lys	Arg	Gln	Lys	Leu	Arg	Cys	Leu	Val
			50				55					60				
	Arg	Gly	Lys	Lys	Glu	Phe	Pro	Val	Arg	Ala	Asp	Lys	Arg	Val	Thr	Phe
			65			70					75			80		
	Asn	Ile	Lys	Gly	Arg	Cys	Asp	Ile	Asp	Gln	Asp	Leu	Asn	Gln	Lys	Asn
45					85				90					95		
	Pro	Lys	Gly	Ala	Ala	Glu	Phe	Ala	Trp	Asn	Ile	Met	Asp	Phe	Lys	Glu
			100					105					110			
	Asp	Gln	Asp	Val	Arg	Ile	Lys	Val	Gly	Tyr	Glu	Met	Phe	Asp	Lys	Val
			115				120						125			
50	Pro	Tyr	Met	Gln	Ile	Arg	Glu	Asn	Asn	Trp	Thr	Leu	Asn	Ala	Asn	Met
			130				135					140				
	Lys	Gly	Lys	Trp	Asn	Leu	Arg	Tyr	Asp	Leu						
			145			150										

## (2) INFORMATION FOR SEQ ID NO:438:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 750 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

(A) NAME/KEY: -

552

(B) LOCATION: 1..750

(D) OTHER INFORMATION: / Ceres Seq. ID 2036457

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:438:

```

5  atgacgactc agatcagcaa gaagagaaag tttgtagcgg acggtgtgtt ctacgctgaa      60
   ttgaatgagg ttcttacaag agagctagca gaggatggtt actctggtgt tgagggttagg      120
   gttactccta tgaggactga gattatcatc agagctacac gtactcagaa tgttctcgg      180
   gagaagggga ggagaattag ggaattgact tcccttggtc agaagagatt caagtttcca      240
   gttgacagtg ttgagcttta tgccgagaag gttaacaaca gaggtctctg tgccatcgct      300
   caggctgagt ctttacgtta caagcttctt ggtggtctcg ctgttcgtag ggcttgctat      360
10  ggtgtgttga ggtttgttat ggagagtgga gctaagggat gcgagggtat cgtgagtgga      420
   aagcttcgtg ctgccagagc caagtctatg aagttcaaag atggctacat ggtgtcatct      480
   ggtcaaccaa ctaaggaata catagactct gcagtgaac atgttttgct tagacaagg      540
   gtgttgggaa tcaagggtgaa ggttatgctt gattgggacc ctaagggcat atcaggacca      600
   aagacaccat tgcctgatgt tgtgatcatt cattctccta aagaagaaga ggccatctat      660
15  gcacctgctc aggttgctgc cccggctgct ctgtagcag atgcaccact cacagccgta      720
   gattaccctg cgatgatccc agtcgcctaa

```

(2) INFORMATION FOR SEQ ID NO:439:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 249 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..249

(D) OTHER INFORMATION: / Ceres Seq. ID 2036458

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:439:

```

30  Met Thr Thr Gln Ile Ser Lys Lys Arg Lys Phe Val Ala Asp Gly Val
   1           5           10           15
   Phe Tyr Ala Glu Leu Asn Glu Val Leu Thr Arg Glu Leu Ala Glu Asp
           20           25           30
   Gly Tyr Ser Gly Val Glu Val Arg Val Thr Pro Met Arg Thr Glu Ile
           35           40           45
35  Ile Ile Arg Ala Thr Arg Thr Gln Asn Val Leu Gly Glu Lys Gly Arg
   50           55           60
   Arg Ile Arg Glu Leu Thr Ser Leu Val Gln Lys Arg Phe Lys Phe Pro
   65           70           75           80
   Val Asp Ser Val Glu Leu Tyr Ala Glu Lys Val Asn Asn Arg Gly Leu
40  Cys Ala Ile Ala Gln Ala Glu Ser Leu Arg Tyr Lys Leu Leu Gly Gly
           85           90           95
   Leu Ala Val Arg Arg Ala Cys Tyr Gly Val Leu Arg Phe Val Met Glu
           100          105          110
   Ser Gly Ala Lys Gly Cys Glu Val Ile Val Ser Gly Lys Leu Arg Ala
45  Ala Arg Ala Lys Ser Met Lys Phe Lys Asp Gly Tyr Met Val Ser Ser
   130          135          140          145
   Gly Gln Pro Thr Lys Glu Tyr Ile Asp Ser Ala Val Arg His Val Leu
50  Leu Arg Gln Gly Val Leu Gly Ile Lys Val Lys Val Met Leu Asp Trp
           150          155          160
   Asp Pro Lys Gly Ile Ser Gly Pro Lys Thr Pro Leu Pro Asp Val Val
           165          170          175
   Ile Ile His Ser Pro Lys Glu Glu Glu Ala Ile Tyr Ala Pro Ala Gln
55  Val Ala Ala Pro Ala Ala Leu Val Ala Asp Ala Pro Leu Thr Ala Val
   180          185          190          195
   Asp Tyr Pro Ala Met Ile Pro Val Ala
           200          205          210
   210          215          220          225
   225          230          235          240
   245

```

(2) INFORMATION FOR SEQ ID NO:440:

(i) SEQUENCE CHARACTERISTICS:

553

(A) LENGTH: 206 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..206

(D) OTHER INFORMATION: / Ceres Seq. ID 2036459

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:440:

```

Met Arg Thr Glu Ile Ile Arg Ala Thr Arg Thr Gln Asn Val Leu
1      5      10      15
Gly Glu Lys Gly Arg Arg Ile Arg Glu Leu Thr Ser Leu Val Gln Lys
      20      25      30
15 Arg Phe Lys Phe Pro Val Asp Ser Val Glu Leu Tyr Ala Glu Lys Val
      35      40      45
Asn Asn Arg Gly Leu Cys Ala Ile Ala Gln Ala Glu Ser Leu Arg Tyr
      50      55      60
Lys Leu Leu Gly Gly Leu Ala Val Arg Arg Ala Cys Tyr Gly Val Leu
20 65      70      75      80
Arg Phe Val Met Glu Ser Gly Ala Lys Gly Cys Glu Val Ile Val Ser
      85      90      95
Gly Lys Leu Arg Ala Ala Arg Ala Lys Ser Met Lys Phe Lys Asp Gly
      100     105     110
25 Tyr Met Val Ser Ser Gly Gln Pro Thr Lys Glu Tyr Ile Asp Ser Ala
      115     120     125
Val Arg His Val Leu Leu Arg Gln Gly Val Leu Gly Ile Lys Val Lys
      130     135     140
Val Met Leu Asp Trp Asp Pro Lys Gly Ile Ser Gly Pro Lys Thr Pro
30 145     150     155     160
Leu Pro Asp Val Val Ile Ile His Ser Pro Lys Glu Glu Glu Ala Ile
      165     170     175
Tyr Ala Pro Ala Gln Val Ala Ala Pro Ala Ala Leu Val Ala Asp Ala
      180     185     190
35 Pro Leu Thr Ala Val Asp Tyr Pro Ala Met Ile Pro Val Ala
      195     200     205

```

(2) INFORMATION FOR SEQ ID NO:441:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 549 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..549

(D) OTHER INFORMATION: / Ceres Seq. ID 2036585

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:441:

```

atggttctca agacggagct ttgtcgtttc agtggacaga agatttacct aggaagagga      60
attaggttta tccgatctga ttctcaggtt ttcttgtttc ttaactcaaa atgtaagagg      120
tacttccata acaagttgaa gccatccaag cttgcatgga ctgccatgta cagaaagcaa      180
cacaagaagg atgcagcaca agaggctgtg aagagaagga gacgtgccac caagaagcca      240
tactcaaggt ccattgttgg tgctaccttg gaagtaattc agaagaagag agctgagaag      300
cctgaagttc gtgatgcagc cagggaagct gctctgcgtg agatcaagga aagaatcaaa      360
55 aagaccaaag atgaaaagaa ggctaagaag gtggaatttg cttctaagca acagaagggtc      420
aaggctaatt tcccaaagc tgctgctgca tccaagggtc ctaagggtgt ggtggcaaac      480
gctgaagagc ttaaagccat cttttctcac tctgcgtctt ttctgctagt agctactttt      540
agtagttga

```

(2) INFORMATION FOR SEQ ID NO:442:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 182 amino acids

(B) TYPE: amino acid

554

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..182

(D) OTHER INFORMATION: / Ceres Seq. ID 2036586

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:442:

5 Met Val Leu Lys Thr Glu Leu Cys Arg Phe Ser Gly Gln Lys Ile Tyr  
 10 1 5 10 15  
 Pro Gly Arg Gly Ile Arg Phe Ile Arg Ser Asp Ser Gln Val Phe Leu  
 20 20 25 30  
 Phe Leu Asn Ser Lys Cys Lys Arg Tyr Phe His Asn Lys Leu Lys Pro  
 35 40 45  
 15 Ser Lys Leu Ala Trp Thr Ala Met Tyr Arg Lys Gln His Lys Lys Asp  
 50 55 60  
 Ala Ala Gln Glu Ala Val Lys Arg Arg Arg Arg Ala Thr Lys Lys Pro  
 65 70 75 80  
 Tyr Ser Arg Ser Ile Val Gly Ala Thr Leu Glu Val Ile Gln Lys Lys  
 20 85 90 95  
 Arg Ala Glu Lys Pro Glu Val Arg Asp Ala Ala Arg Glu Ala Ala Leu  
 100 105 110  
 Arg Glu Ile Lys Glu Arg Ile Lys Lys Thr Lys Asp Glu Lys Lys Ala  
 115 120 125  
 25 Lys Lys Val Glu Phe Ala Ser Lys Gln Gln Lys Val Lys Ala Asn Phe  
 130 135 140  
 Pro Lys Ala Ala Ala Ala Ser Lys Gly Pro Lys Val Val Val Ala Asn  
 145 150 155 160  
 Ala Glu Glu Leu Lys Ala Ile Phe Ser His Ser Ala Ser Phe Leu Leu  
 30 165 170 175  
 Val Ala Thr Phe Ser Ser  
 180

(2) INFORMATION FOR SEQ ID NO:443:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 127 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..127

(D) OTHER INFORMATION: / Ceres Seq. ID 2036587

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:443:

45 Met Tyr Arg Lys Gln His Lys Lys Asp Ala Ala Gln Glu Ala Val Lys  
 1 5 10 15  
 Arg Arg Arg Arg Ala Thr Lys Lys Pro Tyr Ser Arg Ser Ile Val Gly  
 20 25 30  
 Ala Thr Leu Glu Val Ile Gln Lys Lys Arg Ala Glu Lys Pro Glu Val  
 35 40 45  
 50 Arg Asp Ala Ala Arg Glu Ala Ala Leu Arg Glu Ile Lys Glu Arg Ile  
 50 55 60  
 Lys Lys Thr Lys Asp Glu Lys Lys Ala Lys Lys Val Glu Phe Ala Ser  
 65 70 75 80  
 55 Lys Gln Gln Lys Val Lys Ala Asn Phe Pro Lys Ala Ala Ala Ala Ser  
 85 90 95  
 Lys Gly Pro Lys Val Val Val Ala Asn Ala Glu Glu Leu Lys Ala Ile  
 100 105 110  
 Phe Ser His Ser Ala Ser Phe Leu Val Ala Thr Phe Ser Ser  
 60 115 120 125

(2) INFORMATION FOR SEQ ID NO:444:

(i) SEQUENCE CHARACTERISTICS:

555

- (A) LENGTH: 486 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear
- 5 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..486  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2039554
- 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:444:  
 atggctgtcg tcggcgctcc aatatcgtct ccggcggtc agctgcagac acaattttctc 60  
 tccaatccca ttctcccccg ctttcgccgg tctttctcca ccggaaaatc accagcaact 120  
 ttctccgtcg tagctatggc tccccagaaa aaggtgaaca aatatgatgc caagtggaag 180  
 aaacaatggt acggagctgg attgtttttc gaagggagtg agcaaataaa cgttgatgtt 240  
 15 ttcaagaagc tggagaagcg aaaagtgttg agcaacgttg agaaatctgg cctgctgtca 300  
 aaagcagagg ggttgggact cacattgtca tctcttgaga agcttaaagt cttctccaaa 360  
 gcagaggacc ttggtcttct cagtctcctt gagaacttag ctggaacatc gcctgcggtc 420  
 ttagcctcgg ctgcattacc agctctcacg agctgctatt gtagccgtgg tgttgatccc 480  
 ggatga
- 20 (2) INFORMATION FOR SEQ ID NO:445:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 161 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear
- 25 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..161  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2039555
- 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:445:  
 Met Ala Val Val Gly Ala Pro Ile Ser Ser Pro Ala Ala Gln Leu Gln  
 1 5 10 15  
 Thr Gln Phe Leu Ser Asn Pro Ile Leu Pro Arg Phe Arg Arg Ser Phe  
 35 20 25 30  
 Ser Thr Gly Lys Ser Pro Ala Thr Phe Ser Val Val Ala Met Ala Pro  
 35 40 45  
 Gln Lys Lys Val Asn Lys Tyr Asp Ala Lys Trp Lys Lys Gln Trp Tyr  
 50 55 60  
 40 Gly Ala Gly Leu Phe Phe Glu Gly Ser Glu Gln Ile Asn Val Asp Val  
 65 70 75 80  
 Phe Lys Lys Leu Glu Lys Arg Lys Val Leu Ser Asn Val Glu Lys Ser  
 85 90 95  
 Gly Leu Leu Ser Lys Ala Glu Gly Leu Gly Leu Thr Leu Ser Ser Leu  
 45 100 105 110  
 Glu Lys Leu Lys Val Phe Ser Lys Ala Glu Asp Leu Gly Leu Leu Ser  
 115 120 125  
 Leu Leu Glu Asn Leu Ala Gly Thr Ser Pro Ala Val Leu Ala Ser Ala  
 130 135 140  
 50 Ala Leu Pro Ala Leu Thr Ser Cys Tyr Cys Ser Arg Gly Val Asp Pro  
 145 150 155 160  
 Gly
- 55 (2) INFORMATION FOR SEQ ID NO:446:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 116 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear
- 60 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide

556

(B) LOCATION: 1..116

(D) OTHER INFORMATION: / Ceres Seq. ID 2039556

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:446:

5 Met Ala Pro Gln Lys Lys Val Asn Lys Tyr Asp Ala Lys Trp Lys Lys  
 1 5 10 15  
 Gln Trp Tyr Gly Ala Gly Leu Phe Phe Glu Gly Ser Glu Gln Ile Asn  
 20 25 30  
 Val Asp Val Phe Lys Lys Leu Glu Lys Arg Lys Val Leu Ser Asn Val  
 35 40 45  
 10 Glu Lys Ser Gly Leu Leu Ser Lys Ala Glu Gly Leu Gly Leu Thr Leu  
 50 55 60  
 Ser Ser Leu Glu Lys Leu Lys Val Phe Ser Lys Ala Glu Asp Leu Gly  
 65 70 75 80  
 Leu Leu Ser Leu Leu Glu Asn Leu Ala Gly Thr Ser Pro Ala Val Leu  
 85 90 95  
 15 Ala Ser Ala Ala Leu Pro Ala Leu Thr Ser Cys Tyr Cys Ser Arg Gly  
 100 105 110  
 Val Asp Pro Gly  
 115

20 (2) INFORMATION FOR SEQ ID NO:447:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 378 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

25 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..378

30 (D) OTHER INFORMATION: / Ceres Seq. ID 2044283

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:447:

atggcggaatt tgatgatgag attaccaatt agcttgagaa gcttctctgt ttcagcttct 60  
 tcatccaacg gttcgccgcc ggtgatcgga ggatctagcg gcggtgtagg accgatgatt 120  
 gtggaattac cggttgagaa gatacgaaga ccgttgatgc gaaccagatc caacgatcag 180  
 35 aacaaagtga aagagcttat ggatagtatc cgtcaaactc gtcttcaagt tccgattgat 240  
 gtgattgaag ttgatggaac ttactatggg ttctcgggat gtcacagata cgaggcgcat 300  
 cagaagctag ggcttccaac tatacgttgc aaaatccgta aaggaacaaa ggaaacatta 360  
 aggcacatc ttcgctga

40 (2) INFORMATION FOR SEQ ID NO:448:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 125 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

45 (ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..125

(D) OTHER INFORMATION: / Ceres Seq. ID 2044284

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:448:

Met Ala Asn Leu Met Met Arg Leu Pro Ile Ser Leu Arg Ser Phe Ser  
 1 5 10 15  
 Val Ser Ala Ser Ser Ser Asn Gly Ser Pro Pro Val Ile Gly Gly Ser  
 20 25 30  
 55 Ser Gly Gly Val Gly Pro Met Ile Val Glu Leu Pro Leu Glu Lys Ile  
 35 40 45  
 Arg Arg Pro Leu Met Arg Thr Arg Ser Asn Asp Gln Asn Lys Val Lys  
 50 55 60  
 Glu Leu Met Asp Ser Ile Arg Gln Ile Gly Leu Gln Val Pro Ile Asp  
 65 70 75 80  
 60 Val Ile Glu Val Asp Gly Thr Tyr Tyr Gly Phe Ser Gly Cys His Arg  
 85 90 95



557

Tyr Glu Ala His Gln Lys Leu Gly Leu Pro Thr Ile Arg Cys Lys Ile  
 100 105 110  
 Arg Lys Gly Thr Lys Glu Thr Leu Arg His His Leu Arg  
 115 120 125

## 5 (2) INFORMATION FOR SEQ ID NO:449:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 121 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

## 10 (ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..121

## 15 (D) OTHER INFORMATION: / Ceres Seq. ID 2044285

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:449:

Met Met Arg Leu Pro Ile Ser Leu Arg Ser Phe Ser Val Ser Ala Ser  
 1 5 10 15  
 Ser Ser Asn Gly Ser Pro Pro Val Ile Gly Gly Ser Ser Gly Gly Val  
 20 20 25 30  
 Gly Pro Met Ile Val Glu Leu Pro Leu Glu Lys Ile Arg Arg Pro Leu  
 35 40 45  
 Met Arg Thr Arg Ser Asn Asp Gln Asn Lys Val Lys Glu Leu Met Asp  
 50 55 60  
 Ser Ile Arg Gln Ile Gly Leu Gln Val Pro Ile Asp Val Ile Glu Val  
 25 65 70 75 80  
 Asp Gly Thr Tyr Tyr Gly Phe Ser Gly Cys His Arg Tyr Glu Ala His  
 85 90 95  
 Gln Lys Leu Gly Leu Pro Thr Ile Arg Cys Lys Ile Arg Lys Gly Thr  
 30 100 105 110  
 Lys Glu Thr Leu Arg His His Leu Arg  
 115 120

## (2) INFORMATION FOR SEQ ID NO:450:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 120 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..120

(D) OTHER INFORMATION: / Ceres Seq. ID 2044286

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:450:

Met Arg Leu Pro Ile Ser Leu Arg Ser Phe Ser Val Ser Ala Ser Ser  
 1 5 10 15  
 Ser Asn Gly Ser Pro Pro Val Ile Gly Gly Ser Ser Gly Gly Val Gly  
 20 25 30  
 Pro Met Ile Val Glu Leu Pro Leu Glu Lys Ile Arg Arg Pro Leu Met  
 35 40 45  
 Arg Thr Arg Ser Asn Asp Gln Asn Lys Val Lys Glu Leu Met Asp Ser  
 50 55 60  
 Ile Arg Gln Ile Gly Leu Gln Val Pro Ile Asp Val Ile Glu Val Asp  
 65 70 75 80  
 Gly Thr Tyr Tyr Gly Phe Ser Gly Cys His Arg Tyr Glu Ala His Gln  
 85 90 95  
 Lys Leu Gly Leu Pro Thr Ile Arg Cys Lys Ile Arg Lys Gly Thr Lys  
 100 105 110  
 Glu Thr Leu Arg His His Leu Arg  
 115 120

## (2) INFORMATION FOR SEQ ID NO:451:

## (i) SEQUENCE CHARACTERISTICS:

558

(A) LENGTH: 1236 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..1236

(D) OTHER INFORMATION: / Ceres Seq. ID 2048114

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:451:

```

atgagaagac ctagtcaa atgaggctt ctattaacat cctttttcgg tgttattggt      60
ggtttcctta tgggtattac ttttccaacc ttgactttta ctaagatgaa tcttccatcc      120
acattgtttc cctcgattga tcttgcatatc attgaggata aatactctga catatcaaga      180
caaagactat ttggttcttg gtcttcgaca aaaggcctca aactcaagaa tgacatccct      240
15 gaccctccat ataactataa tgacactaag gttgatgata gaacgttcga gctattgcag      300
atatgggttt cgactaacc cgtggtgct gagaggctac caccagatat agtcacgcct      360
gaatcagatt ttacctccg tcgactgtgg ggcgacccta atgaggattt aacagtcaag      420
cagcgggtatc tagtaacatt tacggttggc tatgatcaga ggaaaaatat agacactgtg      480
ttgaagaagt tctcagataa cttctctata atgctgtttc actacgatgg cggggcaagc      540
20 gaatgggaag agtttgaatg gtccaagcga gccattcatg tgagcattcg gaaacaaaca      600
aaatggtggt acgctaagcg atttcttcat cctgacatag ttgccccta tgaatatatc      660
ttcatatggg atgaggatct tggcggtgaa cactttgatt cgaaaaata tctggcggtg      720
gtgaagaagc atggtttggc aatctcacag cctggattag agccatatga agggctcaca      780
tgaggagatga ccaagaaaag agacgacact gaagtccaca agcatgctga ggaaaggaat      840
25 ggggtggtgca ctgatcccaa tttaccccct tgtgcagcgt ttgtggagat tatggctcct      900
gttttctccc gcaaggcatg gcgctgtgtg tggcatatga ttcagaacga tttgattcat      960
ggatggggtc tggactttgc cgttcggaaa tgtgttcaga acgcacacga gaaaattgga      1020
gttgtagatg ctcaatggat tatacatcaa ggtgttccat cattagggaa tcaaggacaa      1080
30 ccagagcaag ggaaacaacc atgggaaggg gtgagagaac gatgcaggag agagtggaca      1140
atgtttcaag acagattgga tgatgctgaa aaagcttatt ttgaagcatc tgctcacaag      1200
aatgcttctt cagggcctca cgggaattgg gtatag

```

(2) INFORMATION FOR SEQ ID NO:452:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 411 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..411

(D) OTHER INFORMATION: / Ceres Seq. ID 2048115

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:452:

```

Met Arg Arg Pro Ser Gln Met Met Arg Leu Leu Leu Thr Ser Phe Phe
45 1      5      10      15
Gly Val Ile Val Gly Phe Leu Met Gly Ile Thr Phe Pro Thr Leu Thr
      20      25      30
Leu Thr Lys Met Asn Leu Pro Ser Thr Leu Phe Pro Ser Ile Asp Leu
      35      40      45
50 Ala Tyr Ile Glu Asp Lys Tyr Ser Asp Ile Ser Arg Gln Arg Leu Phe
      50      55      60
Gly Ser Trp Ser Ser Thr Lys Gly Leu Lys Leu Lys Asn Asp Ile Pro
65      70      75      80
Asp Pro Pro Tyr Asn Tyr Asn Asp Thr Lys Val Asp Asp Arg Thr Phe
      85      90      95
55 Glu Leu Leu Gln Ile Trp Val Ser Thr Asn Pro Arg Gly Ala Glu Arg
      100      105      110
Leu Pro Pro Asp Ile Val Thr Pro Glu Ser Asp Phe Tyr Leu Arg Arg
      115      120      125
60 Leu Trp Gly Asp Pro Asn Glu Asp Leu Thr Val Lys Gln Arg Tyr Leu
      130      135      140
Val Thr Phe Thr Val Gly Tyr Asp Gln Arg Lys Asn Ile Asp Thr Val

```

559

	145				150					155					160	
	Leu	Lys	Lys	Phe	Ser	Asp	Asn	Phe	Ser	Ile	Met	Leu	Phe	His	Tyr	Asp
					165					170					175	
5	Gly	Arg	Ala	Ser	Glu	Trp	Glu	Glu	Phe	Glu	Trp	Ser	Lys	Arg	Ala	Ile
				180					185					190		
	His	Val	Ser	Ile	Arg	Lys	Gln	Thr	Lys	Trp	Trp	Tyr	Ala	Lys	Arg	Phe
			195				200					205				
	Leu	His	Pro	Asp	Ile	Val	Ala	Pro	Tyr	Glu	Tyr	Ile	Phe	Ile	Trp	Asp
		210				215					220					
10	Glu	Asp	Leu	Gly	Val	Glu	His	Phe	Asp	Ser	Glu	Lys	Tyr	Leu	Ala	Val
	225					230					235				240	
	Val	Lys	Lys	His	Gly	Leu	Glu	Ile	Ser	Gln	Pro	Gly	Leu	Glu	Pro	Tyr
				245					250					255		
	Glu	Gly	Leu	Thr	Trp	Glu	Met	Thr	Lys	Lys	Arg	Asp	Asp	Thr	Glu	Val
15				260					265					270		
	His	Lys	His	Ala	Glu	Glu	Arg	Asn	Gly	Trp	Cys	Thr	Asp	Pro	Asn	Leu
			275					280				285				
	Pro	Pro	Cys	Ala	Ala	Phe	Val	Glu	Ile	Met	Ala	Pro	Val	Phe	Ser	Arg
		290				295					300					
20	Lys	Ala	Trp	Arg	Cys	Val	Trp	His	Met	Ile	Gln	Asn	Asp	Leu	Ile	His
	305					310					315				320	
	Gly	Trp	Gly	Leu	Asp	Phe	Ala	Val	Arg	Lys	Cys	Val	Gln	Asn	Ala	His
				325					330					335		
	Glu	Lys	Ile	Gly	Val	Val	Asp	Ala	Gln	Trp	Ile	Ile	His	Gln	Gly	Val
25				340					345					350		
	Pro	Ser	Leu	Gly	Asn	Gln	Gly	Gln	Pro	Glu	Gln	Gly	Lys	Gln	Pro	Trp
			355				360					365				
	Glu	Gly	Val	Arg	Glu	Arg	Cys	Arg	Arg	Glu	Trp	Thr	Met	Phe	Gln	Asp
		370				375					380					
30	Arg	Leu	Asp	Asp	Ala	Glu	Lys	Ala	Tyr	Phe	Glu	Ala	Ser	Ala	His	Lys
	385					390					395				400	
	Asn	Ala	Ser	Ser	Arg	Pro	His	Gly	Asn	Trp	Val					
				405					410							

(2) INFORMATION FOR SEQ ID NO:453:

35	(i) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 405 amino acids
	(B) TYPE: amino acid
	(C) STRANDEDNESS:
	(D) TOPOLOGY: linear
40	(ii) MOLECULE TYPE: peptide
	(ix) FEATURE:
	(A) NAME/KEY: peptide
	(B) LOCATION: 1..405
	(D) OTHER INFORMATION: / Ceres Seq. ID 2048116
45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:453:
	Met Met Arg Leu Leu Leu Thr Ser Phe Phe Gly Val Ile Val Gly Phe
	1 5 10 15
	Leu Met Gly Ile Thr Phe Pro Thr Leu Thr Lys Met Asn Leu
	20 25 30
50	Pro Ser Thr Leu Phe Pro Ser Ile Asp Leu Ala Tyr Ile Glu Asp Lys
	35 40 45
	Tyr Ser Asp Ile Ser Arg Gln Arg Leu Phe Gly Ser Trp Ser Ser Thr
	50 55 60
	Lys Gly Leu Lys Leu Lys Asn Asp Ile Pro Asp Pro Pro Tyr Asn Tyr
55	65 70 75 80
	Asn Asp Thr Lys Val Asp Asp Arg Thr Phe Glu Leu Leu Gln Ile Trp
	85 90 95
	Val Ser Thr Asn Pro Arg Gly Ala Glu Arg Leu Pro Pro Asp Ile Val
	100 105 110
60	Thr Pro Glu Ser Asp Phe Tyr Leu Arg Arg Leu Trp Gly Asp Pro Asn
	115 120 125
	Glu Asp Leu Thr Val Lys Gln Arg Tyr Leu Val Thr Phe Thr Val Gly

560

130 135 140  
 Tyr Asp Gln Arg Lys Asn Ile Asp Thr Val Leu Lys Lys Phe Ser Asp  
 145 150 155 160  
 5 Asn Phe Ser Ile Met Leu Phe His Tyr Asp Gly Arg Ala Ser Glu Trp  
 165 170 175  
 Glu Glu Phe Glu Trp Ser Lys Arg Ala Ile His Val Ser Ile Arg Lys  
 180 185 190  
 Gln Thr Lys Trp Trp Tyr Ala Lys Arg Phe Leu His Pro Asp Ile Val  
 195 200 205  
 10 Ala Pro Tyr Glu Tyr Ile Phe Ile Trp Asp Glu Asp Leu Gly Val Glu  
 210 215 220  
 His Phe Asp Ser Glu Lys Tyr Leu Ala Val Val Lys Lys His Gly Leu  
 225 230 235 240  
 15 Glu Ile Ser Gln Pro Gly Leu Glu Pro Tyr Glu Gly Leu Thr Trp Glu  
 245 250 255  
 Met Thr Lys Lys Arg Asp Asp Thr Glu Val His Lys His Ala Glu Glu  
 260 265 270  
 Arg Asn Gly Trp Cys Thr Asp Pro Asn Leu Pro Pro Cys Ala Ala Phe  
 275 280 285  
 20 Val Glu Ile Met Ala Pro Val Phe Ser Arg Lys Ala Trp Arg Cys Val  
 290 295 300  
 Trp His Met Ile Gln Asn Asp Leu Ile His Gly Trp Gly Leu Asp Phe  
 305 310 315 320  
 25 Ala Val Arg Lys Cys Val Gln Asn Ala His Glu Lys Ile Gly Val Val  
 325 330 335  
 Asp Ala Gln Trp Ile Ile His Gln Gly Val Pro Ser Leu Gly Asn Gln  
 340 345 350  
 Gly Gln Pro Glu Gln Gly Lys Gln Pro Trp Glu Gly Val Arg Glu Arg  
 355 360 365  
 30 Cys Arg Arg Glu Trp Thr Met Phe Gln Asp Arg Leu Asp Asp Ala Glu  
 370 375 380  
 Lys Ala Tyr Phe Glu Ala Ser Ala His Lys Asn Ala Ser Ser Arg Pro  
 385 390 395 400  
 35 His Gly Asn Trp Val  
 405

## (2) INFORMATION FOR SEQ ID NO:454:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 404 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..404

(D) OTHER INFORMATION: / Ceres Seq. ID 2048117

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:454:

Met Arg Leu Leu Leu Thr Ser Phe Phe Gly Val Ile Val Gly Phe Leu  
 1 5 10 15  
 50 Met Gly Ile Thr Phe Pro Thr Leu Thr Lys Met Asn Leu Pro  
 20 25 30  
 Ser Thr Leu Phe Pro Ser Ile Asp Leu Ala Tyr Ile Glu Asp Lys Tyr  
 35 40 45  
 55 Ser Asp Ile Ser Arg Gln Arg Leu Phe Gly Ser Trp Ser Ser Thr Lys  
 50 55 60  
 Gly Leu Lys Leu Lys Asn Asp Ile Pro Asp Pro Pro Tyr Asn Tyr Asn  
 65 70 75 80  
 Asp Thr Lys Val Asp Asp Arg Thr Phe Glu Leu Leu Gln Ile Trp Val  
 85 90 95  
 60 Ser Thr Asn Pro Arg Gly Ala Glu Arg Leu Pro Pro Asp Ile Val Thr  
 100 105 110  
 Pro Glu Ser Asp Phe Tyr Leu Arg Arg Leu Trp Gly Asp Pro Asn Glu

561

[illegible]

(2) INFORMATION FOR SEQ ID NO:455:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 649 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..649

(D) OTHER INFORMATION: / Ceres Seq. ID 2048271

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:455:

50	ggaatttctc	actctctcta	tctctcttag	ccagctctga	tcatttcgat	ttttttccgg	60
	tgaaaaggga	ggagaaacat	ggttgtagct	atcagactgt	cgagatttgg	atgcaaaaat	120
	cggccatttt	ttagggttat	ggctgctgat	agcagatctc	caagagacgg	gaagcatctt	180
	gaggtcttag	gttacttcaa	tcctttgcc	ggccaggacg	gtggttaagag	gatgggtctc	240
	aagttcgatc	gaattaagta	ctggttatct	gttggtgctc	agccatcaga	cccggttcaa	300
55	cgtctcctat	tcagatccgg	tttacttcct	cctcctccaa	tgggtggctat	gggacgtaaa	360
	ggtggagcac	gagacacacg	cccagttgat	ccgatgactg	gtcgcctatgt	ggatgcagag	420
	aataaaacag	ttaatgccaa	tgataaaccg	cctaaggaa	aggatacaga	agccaagagt	480
	gcatgattca	ttagcctttct	gtcatcgtag	cttttcaagt	tcactttgtt	gtcgattata	540
	ttgtgtaatg	cagcattaga	caactgactt	gttttcctttg	tttggcgata	aacggcaagg	600
60	tttttggcac	tttttgcaga	aacggcacat	attttgcatt	gggatatttt		

(2) INFORMATION FOR SEQ ID NO:456:

(i) SEQUENCE CHARACTERISTICS:

562

(A) LENGTH: 135 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..135

(D) OTHER INFORMATION: / Ceres Seq. ID 2048272

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:456:

Met Val Val Arg Ile Arg Leu Ser Arg Phe Gly Cys Lys Asn Arg Pro  
 1 5 10 15  
 Phe Phe Arg Val Met Ala Ala Asp Ser Arg Ser Pro Arg Asp Gly Lys  
 20 25 30  
 15 His Leu Glu Val Leu Gly Tyr Phe Asn Pro Leu Pro Gly Gln Asp Gly  
 35 40 45  
 Gly Lys Arg Met Gly Leu Lys Phe Asp Arg Ile Lys Tyr Trp Leu Ser  
 50 55 60  
 Val Gly Ala Gln Pro Ser Asp Pro Val Gln Arg Leu Leu Phe Arg Ser  
 65 70 75 80  
 20 Gly Leu Leu Pro Pro Pro Pro Met Val Ala Met Gly Arg Lys Gly Gly  
 85 90 95  
 Ala Arg Asp Thr Arg Pro Val Asp Pro Met Thr Gly Arg Tyr Val Asp  
 100 105 110  
 25 Ala Glu Asn Lys Thr Val Asn Ala Asn Asp Asn Gln Pro Lys Glu Glu  
 115 120 125  
 Asp Thr Glu Ala Lys Ser Ala  
 130 135

(2) INFORMATION FOR SEQ ID NO:457:

30 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 115 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..115

(D) OTHER INFORMATION: / Ceres Seq. ID 2048273

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:457:

Met Ala Ala Asp Ser Arg Ser Pro Arg Asp Gly Lys His Leu Glu Val  
 1 5 10 15  
 Leu Gly Tyr Phe Asn Pro Leu Pro Gly Gln Asp Gly Gly Lys Arg Met  
 20 25 30  
 45 Gly Leu Lys Phe Asp Arg Ile Lys Tyr Trp Leu Ser Val Gly Ala Gln  
 35 40 45  
 Pro Ser Asp Pro Val Gln Arg Leu Leu Phe Arg Ser Gly Leu Leu Pro  
 50 55 60  
 Pro Pro Pro Met Val Ala Met Gly Arg Lys Gly Gly Ala Arg Asp Thr  
 65 70 75 80  
 50 Arg Pro Val Asp Pro Met Thr Gly Arg Tyr Val Asp Ala Glu Asn Lys  
 85 90 95  
 Thr Val Asn Ala Asn Asp Asn Gln Pro Lys Glu Glu Asp Thr Glu Ala  
 100 105 110  
 55 Lys Ser Ala  
 115

(2) INFORMATION FOR SEQ ID NO:458:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 84 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

60

563

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..84

(D) OTHER INFORMATION: / Ceres Seq. ID 2048274

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:458:

Met Gly Leu Lys Phe Asp Arg Ile Lys Tyr Trp Leu Ser Val Gly Ala  
 1 5 10 15  
 Gln Pro Ser Asp Pro Val Gln Arg Leu Phe Arg Ser Gly Leu Leu  
 20 25 30  
 Pro Pro Pro Pro Met Val Ala Met Gly Arg Lys Gly Gly Ala Arg Asp  
 35 40 45  
 Thr Arg Pro Val Asp Pro Met Thr Gly Arg Tyr Val Asp Ala Glu Asn  
 50 55 60  
 Lys Thr Val Asn Ala Asn Asp Asn Gln Pro Lys Glu Glu Asp Thr Glu  
 65 70 75 80  
 Ala Lys Ser Ala

(2) INFORMATION FOR SEQ ID NO:459:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 285 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..285

(D) OTHER INFORMATION: / Ceres Seq. ID 2048331

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:459:

atgaaactgt gtaatcacca tactggaaag tttgtttcgg agagcgctcga agatcagact 60  
 gagcaggtac tcaaaaacat gggggagata ttgaaagcta gtggtgctga ttattcctcg 120  
 gtggtgaaga caacaatcat gttggctgat ttggctgact tcaagacagt gaacgagata 180  
 tatgccaaat acttcccagc tccttctcca gcacgatcga cgtatcaagt tgcagctttg 240  
 cctctaaacg ccaagatcga gattgaatgt attgcaacac tctag

(2) INFORMATION FOR SEQ ID NO:460:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 94 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..94

(D) OTHER INFORMATION: / Ceres Seq. ID 2048332

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:460:

Met Lys Leu Cys Asn His His Thr Gly Lys Phe Val Ser Glu Ser Val  
 1 5 10 15  
 Glu Asp Gln Thr Glu Gln Val Leu Lys Asn Met Gly Glu Ile Leu Lys  
 20 25 30  
 Ala Ser Gly Ala Asp Tyr Ser Ser Val Val Lys Thr Thr Ile Met Leu  
 35 40 45  
 Ala Asp Leu Ala Asp Phe Lys Thr Val Asn Glu Ile Tyr Ala Lys Tyr  
 50 55 60  
 Phe Pro Ala Pro Ser Pro Ala Arg Ser Thr Tyr Gln Val Ala Ala Leu  
 65 70 75 80  
 Pro Leu Asn Ala Lys Ile Glu Ile Glu Cys Ile Ala Thr Leu  
 85 90

(2) INFORMATION FOR SEQ ID NO:461:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 68 amino acids

564

(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..68

(D) OTHER INFORMATION: / Ceres Seq. ID 2048333

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:461:

```

10 Met Gly Glu Ile Leu Lys Ala Ser Gly Ala Asp Tyr Ser Ser Val Val
    1           5           10           15
    Lys Thr Thr Ile Met Leu Ala Asp Leu Ala Asp Phe Lys Thr Val Asn
        20           25           30
    Glu Ile Tyr Ala Lys Tyr Phe Pro Ala Pro Ser Pro Ala Arg Ser Thr
15   35           40           45
    Tyr Gln Val Ala Ala Leu Pro Leu Asn Ala Lys Ile Glu Ile Glu Cys
    50           55           60
    Ile Ala Thr Leu
    65

```

(2) INFORMATION FOR SEQ ID NO:462:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 48 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..48

(D) OTHER INFORMATION: / Ceres Seq. ID 2048334

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:462:

```

30 Met Leu Ala Asp Leu Ala Asp Phe Lys Thr Val Asn Glu Ile Tyr Ala
    1           5           10           15
    Lys Tyr Phe Pro Ala Pro Ser Pro Ala Arg Ser Thr Tyr Gln Val Ala
35   20           25           30
    Ala Leu Pro Leu Asn Ala Lys Ile Glu Ile Glu Cys Ile Ala Thr Leu
    35           40           45

```

(2) INFORMATION FOR SEQ ID NO:463:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 627 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..627

(D) OTHER INFORMATION: / Ceres Seq. ID 2048466

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:463:

```

50 atgatggaca agaatgaaaa ccctagtgttt ttgctctctc gaagttccgt gcgactctct      60
    gctcgagcaa gatctcctac cgatacaatg attatctcag aggctaaccg caaagaaatc      120
    tgcaagtacc tcttcaaaga aggagtttgc tttgcgaaga aggatttcaa tcttgctaag      180
    catccgttga tcatgtgcc caacctacaa gtgattaagc ttatgcagag tttcaaatcc      240
    aaggagtatg ttagggagac atttgccctg atgcattact attgggtttt gaccaatgaa      300
55 gggatcgagt tcttgagaac ttatcttaat cttccatctg atgttggtcc tgctactttg      360
    aagaagtcag ctaagcctgg tggtcgtcca tttggtggcc cacctggtga tcgctcaaga      420
    ggacctcgcc atgaaggagg agaccgtccc aggtttggtg accgtgatgg gtaccgtgca      480
    ggtcctcgag cttggtggtg gtttgagggt gaaaagggtg gagctccgc agattaccag      540
    ccatctttcc aagggaagtgg ccgtggtttt ggccgtggtg ctgggtggcta cagcgcagct      600
60 gcaccatctg gttcaggttt gccttga

```

(2) INFORMATION FOR SEQ ID NO:464:

(i) SEQUENCE CHARACTERISTICS:



565

(A) LENGTH: 208 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..208

(D) OTHER INFORMATION: / Ceres Seq. ID 2048467

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:464:

```

Met Met Asp Lys Asn Glu Asn Pro Ser Phe Leu Leu Ser Arg Ser Ser
1      5      10      15
Val Arg Leu Ser Ala Arg Ala Arg Ser Pro Thr Asp Thr Met Ile Ile
      20      25      30
15 Ser Glu Ala Asn Arg Lys Glu Ile Cys Lys Tyr Leu Phe Lys Glu Gly
      35      40      45
Val Cys Phe Ala Lys Lys Asp Phe Asn Leu Ala Lys His Pro Leu Ile
      50      55      60
20 Asp Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln Ser Phe Lys Ser
      65      70      75      80
Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His Tyr Tyr Trp Phe
      85      90      95
Leu Thr Asn Glu Gly Ile Glu Phe Leu Arg Thr Tyr Leu Asn Leu Pro
      100     105     110
25 Ser Asp Val Val Pro Ala Thr Leu Lys Lys Ser Ala Lys Pro Gly Gly
      115     120     125
Arg Pro Phe Gly Gly Pro Pro Gly Asp Arg Ser Arg Gly Pro Arg His
      130     135     140
Glu Gly Gly Asp Arg Pro Arg Phe Gly Asp Arg Asp Gly Tyr Arg Ala
      145     150     155     160
30 Gly Pro Arg Ala Gly Gly Glu Phe Gly Gly Glu Lys Gly Gly Ala Pro
      165     170     175
Ala Asp Tyr Gln Pro Ser Phe Gln Gly Ser Gly Arg Gly Phe Gly Arg
      180     185     190
35 Gly Ala Gly Gly Tyr Ser Ala Ala Ala Pro Ser Gly Ser Gly Leu Pro
      195     200     205

```

(2) INFORMATION FOR SEQ ID NO:465:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 207 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..207

(D) OTHER INFORMATION: / Ceres Seq. ID 2048468

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:465:

```

Met Asp Lys Asn Glu Asn Pro Ser Phe Leu Leu Ser Arg Ser Ser Val
50 1      5      10      15
Arg Leu Ser Ala Arg Ala Arg Ser Pro Thr Asp Thr Met Ile Ile Ser
      20      25      30
Glu Ala Asn Arg Lys Glu Ile Cys Lys Tyr Leu Phe Lys Glu Gly Val
      35      40      45
55 Cys Phe Ala Lys Lys Asp Phe Asn Leu Ala Lys His Pro Leu Ile Asp
      50      55      60
Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln Ser Phe Lys Ser Lys
      65      70      75      80
60 Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His Tyr Tyr Trp Phe Leu
      85      90      95
Thr Asn Glu Gly Ile Glu Phe Leu Arg Thr Tyr Leu Asn Leu Pro Ser
      100     105     110

```

566

Asp Val Val Pro Ala Thr Leu Lys Lys Ser Ala Lys Pro Gly Gly Arg  
 115 120 125  
 Pro Phe Gly Gly Pro Pro Gly Asp Arg Ser Arg Gly Pro Arg His Glu  
 130 135 140  
 5 Gly Gly Asp Arg Pro Arg Phe Gly Asp Arg Asp Gly Tyr Arg Ala Gly  
 145 150 155 160  
 Pro Arg Ala Gly Gly Glu Phe Gly Gly Glu Lys Gly Gly Ala Pro Ala  
 165 170 175  
 10 Asp Tyr Gln Pro Ser Phe Gln Gly Ser Gly Arg Gly Phe Gly Arg Gly  
 180 185 190  
 Ala Gly Gly Tyr Ser Ala Ala Ala Pro Ser Gly Ser Gly Leu Pro  
 195 200 205

## (2) INFORMATION FOR SEQ ID NO:466:

## (i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 179 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## 20 (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..179

(D) OTHER INFORMATION: / Ceres Seq. ID 2048469

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:466:

25 Met Ile Ile Ser Glu Ala Asn Arg Lys Glu Ile Cys Lys Tyr Leu Phe  
 1 5 10 15  
 Lys Glu Gly Val Cys Phe Ala Lys Lys Asp Phe Asn Leu Ala Lys His  
 20 25 30  
 30 Pro Leu Ile Asp Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln Ser  
 35 40 45  
 Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His Tyr  
 50 55 60  
 Tyr Trp Phe Leu Thr Asn Glu Gly Ile Glu Phe Leu Arg Thr Tyr Leu  
 65 70 75 80  
 35 Asn Leu Pro Ser Asp Val Val Pro Ala Thr Leu Lys Lys Ser Ala Lys  
 85 90 95  
 Pro Gly Gly Arg Pro Phe Gly Gly Pro Pro Gly Asp Arg Ser Arg Gly  
 100 105 110  
 40 Pro Arg His Glu Gly Gly Asp Arg Pro Arg Phe Gly Asp Arg Asp Gly  
 115 120 125  
 Tyr Arg Ala Gly Pro Arg Ala Gly Gly Glu Phe Gly Gly Glu Lys Gly  
 130 135 140  
 Gly Ala Pro Ala Asp Tyr Gln Pro Ser Phe Gln Gly Ser Gly Arg Gly  
 145 150 155 160  
 45 Phe Gly Arg Gly Ala Gly Gly Tyr Ser Ala Ala Ala Pro Ser Gly Ser  
 165 170 175  
 Gly Leu Pro

## (2) INFORMATION FOR SEQ ID NO:467:

## (i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 585 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## 55 (ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..585

(D) OTHER INFORMATION: / Ceres Seq. ID 2050485

## 60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:467:

atcttcggaa agtctcattt ctcgatcccc aattcgtgga ttaggggttaa aagaaccatt  
 ttattctcg tcgcgcaaca acaaattccag atcgaaaaag gaagaagaga tcgaaatggc

60  
 120

567

gttgagaagg gtttacagtg aaatcagagg gaagaagggtg acggagcttc caggctatat 180  
 caaatcgact ttttcaatgg agaccgtgaa gacctctgtg aagagaggac tcgataacta 240  
 caacgaaaaa tacattcaga ccagctccgt tgatcctatc cttcataatct gcttctacgg 300  
 catggctttc tcttaccttg tcgctctccc taatgagcgt cgccatcttg agcatcagca 360  
 5 gcatgctaag gagcacggtg gtcattgatc tcgtgggaat cgtttcgatc tcgagatgat 420  
 tttagggggg ttgctgtgaa atctttctct gcttgatggg gacgacgact caagaattgt 480  
 gtcttattgt ttcgttttct tgaattttcc tggataatgt tgacctaaag gaaaaccttt 540  
 ctttcgaatt acactccatg atagtcaata attgaagcat catga

(2) INFORMATION FOR SEQ ID NO:468:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 128 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..128

(D) OTHER INFORMATION: / Ceres Seq. ID 2050486

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:468:

Ser Ser Glu Ser Leu Ile Ser Arg Ser Pro Ile Arg Gly Leu Gly Leu  
 1 5 10 15  
 Lys Glu Pro Phe Leu Phe Ser Ser Arg Asn Asn Lys Ser Arg Ser Lys  
 20 25 30  
 25 Lys Glu Glu Glu Ile Glu Met Ala Leu Arg Arg Val Tyr Ser Glu Ile  
 35 40 45  
 Arg Gly Lys Lys Val Thr Glu Leu Pro Gly Tyr Ile Lys Ser Thr Phe  
 50 55 60  
 Ser Met Glu Thr Val Lys Thr Ser Val Lys Arg Gly Leu Asp Asn Tyr  
 30 65 70 75 80  
 Asn Glu Lys Tyr Ile Gln Thr Ser Ser Val Asp Pro Ile Leu His Ile  
 85 90 95  
 Cys Phe Tyr Gly Met Ala Phe Ser Tyr Leu Val Ala Leu Pro Asn Glu  
 100 105 110  
 35 Arg Arg His Leu Glu His Gln Gln His Ala Lys Glu His Gly Gly His  
 115 120 125

(2) INFORMATION FOR SEQ ID NO:469:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 90 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..90

(D) OTHER INFORMATION: / Ceres Seq. ID 2050487

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:469:

Met Ala Leu Arg Arg Val Tyr Ser Glu Ile Arg Gly Lys Lys Val Thr  
 50 1 5 10 15  
 Glu Leu Pro Gly Tyr Ile Lys Ser Thr Phe Ser Met Glu Thr Val Lys  
 20 25 30  
 Thr Ser Val Lys Arg Gly Leu Asp Asn Tyr Asn Glu Lys Tyr Ile Gln  
 35 40 45  
 55 Thr Ser Ser Val Asp Pro Ile Leu His Ile Cys Phe Tyr Gly Met Ala  
 50 55 60  
 Phe Ser Tyr Leu Val Ala Leu Pro Asn Glu Arg Arg His Leu Glu His  
 65 70 75 80  
 Gln Gln His Ala Lys Glu His Gly Gly His  
 85 90

(2) INFORMATION FOR SEQ ID NO:470:

(i) SEQUENCE CHARACTERISTICS:

568

(A) LENGTH: 516 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..516

(D) OTHER INFORMATION: / Ceres Seq. ID 2050708

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:470:

```

caacccaaac aaaacataaa aaacaagtgg aagcttttaa acgagaggga gagagcaaaa      60
atggcgcacgt cgggaacgta cgtgacggaa gttccgctaa aaggatcggc cgagaaatac      120
tacaagaggt ggaagaacga gaaccatgtc ttccctgatg ctatcggcca ccacatccaa      180
aatgttaccg ttcacgaagg cgaacatgac tctcacgggt ctatcaggag ttggaactac      240
15 acatgggatg gaaaggagga ggtgttcaag gagagaagag agatagacga tgagaccaa      300
acgttgacgt taagaggact tgagggtcac gtgatggagc agctcaaagt gtacgacgtc      360
gtctaccaat tcattcccaa atctgaggat acctgcatcg gcaaaatcac tttaatatgg      420
gagaagcgca acgatgattc cccagaacca agcgggtaca tgaaattcgt caagagcttg      480
gttgctgaca tgggaaacca cgtagcaaaa acttaa

```

20 (2) INFORMATION FOR SEQ ID NO:471:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 151 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

25 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..151

30 (D) OTHER INFORMATION: / Ceres Seq. ID 2050709

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:471:

```

Met Ala Thr Ser Gly Thr Tyr Val Thr Glu Val Pro Leu Lys Gly Ser
1      5      10      15
Ala Glu Lys Tyr Lys Arg Trp Lys Asn Glu Asn His Val Phe Pro
35 20      25      30
Asp Ala Ile Gly His His Ile Gln Asn Val Thr Val His Glu Gly Glu
35 35      40      45
His Asp Ser His Gly Ser Ile Arg Ser Trp Asn Tyr Thr Trp Asp Gly
50 50      55      60
40 Lys Glu Glu Val Phe Lys Glu Arg Arg Glu Ile Asp Asp Glu Thr Lys
65 70      75      80
Thr Leu Thr Leu Arg Gly Leu Glu Gly His Val Met Glu Gln Leu Lys
85 90      95
45 Val Tyr Asp Val Val Tyr Gln Phe Ile Pro Lys Ser Glu Asp Thr Cys
100 105      110
Ile Gly Lys Ile Thr Leu Ile Trp Glu Lys Arg Asn Asp Asp Ser Pro
115 120      125
Glu Pro Ser Gly Tyr Met Lys Phe Val Lys Ser Leu Val Ala Asp Met
130 135      140
50 Gly Asn His Val Ser Lys Thr
145 150

```

(2) INFORMATION FOR SEQ ID NO:472:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1730 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..1730

(D) OTHER INFORMATION: / Ceres Seq. ID 2050901

569

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:472:

	atcgcgtggt	actcctttct	aaaggaacca	ttgtatatca	tgggaaggctt	gacttacttg	60
	aagcattctt	gctctccaag	ggattcacgg	tcccttctca	gctgaattct	cttgagtacg	120
5	ctatggagat	acttcagaac	atccgtgatc	cctacgaaaa	cgccaacatt	gctcttccag	180
	accactgcc	tgaaagtaaa	aaacagaacc	aaaaacagag	cattgttcga	tataaaaagct	240
	cgagaatcac	cgaaataaag	cttctctcta	gcagggtttg	gaagatcata	taccgtacaa	300
	ggcagttgct	tctaacaaac	atcttagaat	ctcttgtagt	cggtcttgtc	ttaggcacta	360
	tctaccttaa	tatcggaact	ggcaaagaag	gaatcaggaa	acgatttggc	cttttcgcat	420
10	tcacctcac	attctctctc	tcttccacta	ccccaacctt	tccaatattc	attgatgaac	480
	gacctattct	tctccgagaa	acctcaagcg	gactctacag	actctctctc	cacattcttg	540
	caaacacttt	ggttttcttg	ccatacttgc	tactcatcgc	aatcatctac	tctgtctcac	600
	tctatttctt	tgtaggactc	tgttttctat	ggcaagctct	cgcctacttt	gtgctcgtaa	660
	tctggatcat	tgtcctaatt	gctaactctt	ttgtactttt	cttgagctct	ctcgcaccta	720
15	actacattgc	tggAACAtct	tcagtGAcCa	ttcttcttgc	ggctttcttc	ttgttctctg	780
	gttacttcat	ctctaaagag	agtcttccca	agtactggct	cttcatgtac	ttcttctcaa	840
	tgtacaagta	tgcgttggaC	GcacttctGa	tAAAtgAgTa	Ctcgtgtctg	cACAaCAAGT	900
	gcctggctcg	gtttgaggaa	GcttctgtGa	Atagctgctt	agttactgga	ggtgacgtgt	960
	tagacaagaa	tgggcttcat	gagagacaga	ggatatatgt	gttgggacgt	ccgtcattgg	1020
20	tatccggttt	caaaacttgag	aaacaaggaa	tacgtttctt	gagaagcaag	aaaaccctat	1080
	tgctcccctt	tgtctcaagct	cagtgcctct	ttgcagcgaa	agctcaacca	acaagacgat	1140
	ggaatttatcc	gagtattaat	tgtgatcttg	ttcgctcttt	tatcagatta	tctcagagaa	1200
	caatcgcaga	gagatctgca	aatacctttt	caaagctaqt	ggaaccagag	ggagtatgct	1260
	ttgctaagaa	ggatttcaat	ctcccaaagc	atccgttgat	tgatgtacca	aacctgcaag	1320
25	tgattaaagct	catgcagagt	ttcaagtcca	aggagtacgt	tagggagaca	tttgctgga	1380
	tgcattatta	ttggtttctg	actaatgaag	gaattgagtt	cttgagaact	tatcttaacc	1440
	ttccttccga	tgttgtccct	gctactttga	agaagtctgc	taagcccgggt	ggtcgtccct	1500
	ttggtggccc	acctggtgat	cgccaaagag	gaccacctcg	ctctgatgga	gaccgtccca	1560
	gatttggtga	ccgtgtaggA	taccgtggag	gcccacgtgg	tgggtgatga	aagggtggaG	1620
30	ctccagctga	tttccagccg	tctttccaag	gaggtgtgtg	taggcctggt	tttgGCCGTg	1680
	gtgcaggcgg	ttacagtGcA	Gcagcaccat	ctggttcaqq	qttcccttga		

(2) INFORMATION FOR SEQ ID NO:473:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 575 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..575

(D) OTHER INFORMATION: / Ceres Seq. ID 2050902

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:473:

	Arg	Val	Leu	Leu	Leu	Ser	Lys	Gly	Thr	Ile	Val	Tyr	His	Gly	Arg	Leu
45	1				5					10					15	
	Asp	Leu	Leu	Glu	Ala	Phe	Leu	Leu	Ser	Lys	Gly	Phe	Thr	Val	Pro	Ser
			20						25					30		
	Gln	Leu	Asn	Ser	Leu	Glu	Tyr	Ala	Met	Glu	Ile	Leu	Gln	Asn	Ile	Arg
			35					40					45			
50	Asp	Pro	Tyr	Glu	Asn	Ala	Asn	Ile	Ala	Leu	Pro	Asp	His	Cys	Pro	Glu
		50					55					60				
	Ser	Lys	Lys	Gln	Asn	Gln	Lys	Gln	Ser	Ile	Val	Arg	Tyr	Lys	Ser	Ser
	65					70					75					80
	Arg	Ile	Thr	Glu	Ile	Ser	Leu	Leu	Ser	Ser	Arg	Phe	Trp	Lys	Ile	Ile
					85					90					95	
55	Tyr	Arg	Thr	Arg	Gln	Leu	Leu	Leu	Thr	Asn	Ile	Leu	Glu	Ser	Leu	Val
				100					105						110	
	Val	Gly	Leu	Val	Leu	Gly	Thr	Ile	Tyr	Leu	Asn	Ile	Gly	Thr	Gly	Lys
			115					120					125			
60	Glu	Gly	Ile	Arg	Lys	Arg	Phe	Gly	Leu	Phe	Ala	Phe	Thr	Leu	Thr	Phe
		130					135					140				
	Leu	Leu	Ser	Ser	Thr	Thr	Gln	Thr	Leu	Pro	Ile	Phe	Ile	Asp	Glu	Arg
	145					150					155					160

570

	Pro	Ile	Leu	Leu	Arg	Glu	Thr	Ser	Ser	Gly	Leu	Tyr	Arg	Leu	Ser	Ser
					165					170					175	
	His	Ile	Leu	Ala	Asn	Thr	Leu	Val	Phe	Leu	Pro	Tyr	Leu	Leu	Leu	Ile
				180					185					190		
5	Ala	Ile	Ile	Tyr	Ser	Val	Ser	Leu	Tyr	Phe	Leu	Val	Gly	Leu	Cys	Phe
			195					200					205			
	Ser	Trp	Gln	Ala	Leu	Ala	Tyr	Phe	Val	Leu	Val	Ile	Trp	Ile	Ile	Val
		210					215					220				
10	Leu	Met	Ala	Asn	Ser	Phe	Val	Leu	Phe	Leu	Ser	Ser	Leu	Ala	Pro	Asn
	225					230					235				240	
	Tyr	Ile	Ala	Gly	Thr	Ser	Ser	Val	Thr	Ile	Leu	Leu	Ala	Ala	Phe	Phe
					245					250					255	
	Leu	Phe	Ser	Gly	Tyr	Phe	Ile	Ser	Lys	Glu	Ser	Leu	Pro	Lys	Tyr	Trp
			260						265					270		
15	Leu	Phe	Met	Tyr	Phe	Phe	Ser	Met	Tyr	Lys	Tyr	Ala	Leu	Asp	Ala	Leu
		275						280					285			
	Leu	Ile	Asn	Glu	Tyr	Ser	Cys	Leu	His	Asn	Lys	Cys	Leu	Val	Trp	Phe
	290						295					300				
20	Glu	Glu	Ala	Ser	Val	Asn	Ser	Cys	Leu	Val	Thr	Gly	Gly	Asp	Val	Leu
	305					310					315				320	
	Asp	Lys	Asn	Gly	Leu	His	Glu	Arg	Gln	Arg	Ile	Tyr	Val	Leu	Gly	Arg
					325					330					335	
	Pro	Ser	Leu	Val	Ser	Gly	Phe	Lys	Leu	Glu	Lys	Gln	Gly	Ile	Arg	Phe
			340						345					350		
25	Leu	Arg	Ser	Lys	Lys	Thr	Leu	Leu	Pro	Phe	Ala	Gln	Ala	Gln	Cys	
		355						360					365			
	Leu	Phe	Ala	Ala	Lys	Ala	Gln	Pro	Thr	Arg	Arg	Trp	Asn	Tyr	Pro	Ser
	370						375					380				
30	Ile	Asn	Cys	Asp	Leu	Val	Arg	Pro	Phe	Ile	Arg	Leu	Ser	Gln	Arg	Thr
	385					390					395				400	
	Ile	Ala	Glu	Arg	Ser	Ala	Asn	Thr	Phe	Ser	Lys	Leu	Val	Glu	Pro	Glu
					405					410					415	
	Gly	Val	Cys	Phe	Ala	Lys	Lys	Asp	Phe	Asn	Leu	Pro	Lys	His	Pro	Leu
				420					425					430		
35	Ile	Asp	Val	Pro	Asn	Leu	Gln	Val	Ile	Lys	Leu	Met	Gln	Ser	Phe	Lys
		435						440					445			
	Ser	Lys	Glu	Tyr	Val	Arg	Glu	Thr	Phe	Ala	Trp	Met	His	Tyr	Tyr	Trp
		450					455					460				
40	Phe	Leu	Thr	Asn	Glu	Gly	Ile	Glu	Phe	Leu	Arg	Thr	Tyr	Leu	Asn	Leu
	465					470					475				480	
	Pro	Ser	Asp	Val	Val	Pro	Ala	Thr	Leu	Lys	Lys	Ser	Ala	Lys	Pro	Gly
					485					490					495	
	Gly	Arg	Pro	Phe	Gly	Gly	Pro	Pro	Gly	Asp	Arg	Gln	Arg	Gly	Pro	Pro
			500						505					510		
45	Arg	Ser	Asp	Gly	Asp	Arg	Pro	Arg	Phe	Gly	Asp	Arg	Asp	Gly	Tyr	Arg
		515						520					525			
	Gly	Gly	Pro	Arg	Gly	Gly	Asp	Glu	Lys	Gly	Gly	Ala	Pro	Ala	Asp	Phe
	530						535					540				
50	Gln	Pro	Ser	Phe	Gln	Gly	Gly	Gly	Gly	Arg	Pro	Gly	Phe	Gly	Arg	Gly
	545					550					555				560	
	Ala	Gly	Gly	Tyr	Ser	Ala	Ala	Ala	Pro	Ser	Gly	Ser	Gly	Phe	Pro	
					565					570					575	

## (2) INFORMATION FOR SEQ ID NO:474:

## (i) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 535 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## 60 (ix) FEATURE:

- (A) NAME/KEY: peptide  
 (B) LOCATION: 1..535

571

(D) OTHER INFORMATION: / Ceres Seq. ID 2050903

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:474:

1 Met Glu Ile Leu Gln Asn Ile Arg Asp Pro Tyr Glu Asn Ala Asn Ile 15  
5 Ala Leu Pro Asp His Cys Pro Glu Ser Lys Lys Gln Asn Gln Lys Gln 30  
Ser Ile Val Arg Tyr Lys Ser Ser Arg Ile Thr Glu Ile Ser Leu Leu 45  
Ser Ser Arg Phe Thr Lys Ile Ile Tyr Arg Thr Arg Gln Leu Leu 60  
10 Thr Asn Ile Leu Glu Ser Leu Val Val Gly Leu Val Leu Gly Thr Ile 80  
65 Tyr Leu Asn Ile Gly Thr Gly Lys Glu Gly Ile Arg Lys Arg Phe Gly 95  
Leu Phe Ala Phe Thr Leu Thr Phe Leu Ser Ser Thr Thr Gln Thr 110  
Leu Pro Ile Phe Ile Asp Glu Arg Pro Ile Leu Leu Arg Glu Thr Ser 125  
Ser Gly Leu Tyr Arg Leu Ser Ser His Ile Leu Ala Asn Thr Leu Val 140  
Phe Leu Pro Tyr Leu Leu Ile Ala Ile Ile Tyr Ser Val Ser Leu 155  
145 Tyr Phe Leu Val Gly Leu Cys Phe Ser Trp Gln Ala Leu Ala Tyr Phe 170  
25 Val Leu Val Ile Trp Ile Ile Val Leu Met Ala Asn Ser Phe Val Leu 185  
Phe Leu Ser Ser Leu Ala Pro Asn Tyr Ile Ala Gly Thr Ser Ser Val 200  
Thr Ile Leu Leu Ala Ala Phe Phe Leu Phe Ser Gly Tyr Phe Ile Ser 215  
30 Lys Glu Ser Leu Pro Lys Tyr Trp Leu Phe Met Tyr Phe Phe Ser Met 230  
Tyr Lys Tyr Ala Leu Asn Ala Leu Leu Ile Asn Glu Tyr Ser Cys Leu 245  
35 His Asn Lys Cys Leu Val Trp Phe Glu Glu Ala Ser Val Asn Ser Cys 260  
Leu Val Thr Gly Gly Asp Val Leu Asp Lys Asn Gly Leu His Glu Arg 275  
Gln Arg Ile Tyr Val Leu Gly Arg Pro Ser Leu Val Ser Gly Phe Lys 290  
Leu Glu Lys Gln Gly Ile Arg Phe Leu Arg Ser Lys Lys Thr Leu Leu 305  
Leu Pro Phe Ala Gln Ala Gln Cys Leu Phe Ala Ala Lys Ala Gln Pro 320  
40 Thr Arg Arg Trp Asn Tyr Pro Ser Ile Asn Cys Asp Leu Val Arg Pro 335  
Phe Ile Arg Leu Ser Gln Arg Thr Ile Ala Glu Arg Ser Ala Asn Thr 350  
Phe Ser Lys Leu Val Glu Pro Glu Gly Val Cys Phe Ala Lys Lys Asp 365  
50 Phe Asn Leu Pro Lys His Pro Leu Ile Asp Val Pro Asn Leu Gln Val 380  
385 Ile Lys Leu Met Gln Ser Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr 400  
Phe Ala Trp Met His Tyr Trp Phe Leu Thr Asn Glu Gly Ile Glu 415  
Phe Leu Arg Thr Tyr Leu Asn Leu Pro Ser Asp Val Val Pro Ala Thr 430  
Leu Lys Lys Ser Ala Lys Pro Gly Gly Arg Pro Phe Gly Gly Pro Pro 445  
60 Gly Asp Arg Gln Arg Gly Pro Pro Arg Ser Asp Gly Asp Arg 460  
465

572

(D) OTHER INFORMATION: / Ceres Seq. ID 2050904

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:475:

1 Phe Gly Asp Arg Asp Gly Tyr Arg Gly Pro Arg Gly Gly Asp Glu 485  
Lys Gly Gly Ala Pro Ala Asp Phe Gln Pro Ser Phe Gln Gly Gly 495  
5 Gly Arg Pro Gly Phe Gly Arg Gly Ala Gly Gly Tyr Ser Ala Ala 510  
Pro Ser Gly Ser Gly Phe Pro 525  
330  
10 (2) INFORMATION FOR SEQ ID NO:475:  
(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 350 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..350  
(D) OTHER INFORMATION: / Ceres Seq. ID 2050904  
Phe Ala Asn Ser Phe Val Leu Phe Leu Ser Ser Leu Ala Pro Asn Tyr 1  
10  
15 Phe Ala Gly Thr Ser Ser Val Thr Ile Leu Leu Ala Ala Phe Leu 30  
Phe Ser Gly Tyr Phe Ile Ser Lys Glu Ser Leu Pro Lys Tyr Trp Leu 45  
Phe Met Tyr Phe Phe Ser Met Tyr Lys Tyr Ala Leu Asp Ala Leu Leu 60  
Ile Asn Glu Tyr Ser Cys Leu His Asn Lys Cys Leu Val Trp Phe Glu 75  
80  
30 Glu Ala Ser Val Asn Ser Cys Leu Val Thr Gly Gly Asp Val Leu Asp 95  
Lys Asn Gly Leu His Glu Arg Gln Arg Ile Tyr Val Leu Gly Arg Pro 110  
35 Ser Leu Val Ser Gly Phe Lys Leu Glu Lys Gln Gly Ile Arg Phe Leu 125  
Arg Ser Lys Lys Thr Leu Leu Leu Pro Phe Ala Gln Ala Gln Cys Leu 140  
Phe Ala Ala Lys Ala Gln Pro Thr Arg Arg Trp Asn Tyr Pro Ser Ile 155  
160  
40 Asn Cys Asp Leu Val Arg Pro Phe Ile Arg Leu Ser Gln Arg Thr Ile 175  
Ala Glu Arg Ser Ala Asn Thr Phe Ser Lys Leu Val Glu Pro Glu Gly 190  
Val Cys Phe Ala Lys Lys Asp Phe Asn Leu Pro Lys His Pro Leu Ile 205  
Asp Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln Ser Phe Lys Ser 220  
Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His Tyr Trp Phe 235  
50 Leu Thr Asn Glu Gly Ile Glu Phe Leu Arg Thr Tyr Leu Asn Leu Pro 250  
Ser Asp Val Val Pro Ala Thr Leu Lys Lys Ser Ala Lys Pro Gly Gly 265  
Arg Pro Phe Gly Gly Pro Pro Gly Asp Arg Gln Arg Gly Pro Pro Arg 280  
Ser Asp Gly Asp Arg Pro Arg Phe Gly Asp Arg Asp Gly Tyr Arg Gly 295  
Gly Pro Arg Gly Gly Asp Glu Lys Gly Gly Ala Pro Ala Asp Phe Gln 310  
320  
60 Pro Ser Phe Gln Gly Gly Gly Arg Pro Gly Phe Gly Arg Gly Ala 330  
335

573

Gly Gly Tyr Ser Ala Ala Pro Ser Gly Ser Gly Phe Pro  
340 345 350

(2) INFORMATION FOR SEQ ID NO:476:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 535 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..535  
(D) OTHER INFORMATION: / Ceres Seq. ID 2051325

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:476:

15 gctacataac tcaagatgac aaacatacac atataacata taagaagacc gaatacaaaa  
gatgggagtg agtaggtgtc caaagacacc aactacacatt gctttatcc ttgtacacaa  
tattctcttc ctcaacacct ttaccttgag ttgcgcagac aatactgcc caagagacgt  
tctcaaacct tgcacatgct cgaatgttct caactctacc aattgaagc tgggggcacc  
agctatgagg cctgtgtgtc ctattctctt tggctcaatt gatctcgatg ttctgtttg  
cctttgcacc gcgtccaagc tgcgctctct tggatcaccc atgcacatc ctcttcacct  
taacttggtt cttaagacct ggggggtrac ccttcctcgt ggtattcgtt gccacacata  
gtctacatat attcatattt ttatctatg ttcttgatt ggtttgtct aacatccaaa  
tttcaacct aaattcatgt ataaagaaat aacaaagaaa attaagactt ttatg  
400

(2) INFORMATION FOR SEQ ID NO:477:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 117 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..117  
(D) OTHER INFORMATION: / Ceres Seq. ID 2051326

35 Met Ser Met Ala Pro Lys Thr Ser Thr Thr Leu Ala Leu Phe Leu Val  
1 5 10 15

Thr Asn Ile Leu Phe Leu Asn Leu Ile Thr Leu Ser Cys Ala Asp Asn  
20 25 30 35

Thr Cys Pro Arg Asp Val Leu Lys Leu Ser Thr Cys Ser Asn Val Leu  
40 45

Asn Leu Ile Asn Leu Lys Leu Gly Ala Pro Ala Met Arg Pro Cys Cys  
50 55 60

Ser Ile Leu Phe Gly Leu Ile Asp Leu Asp Val Ala Val Cys Leu Cys  
65 70 75 80

Thr Ala Leu Lys Leu Ser Leu Leu Gly Ile Thr Ile Asp Thr Pro Ile  
85 90 95

His Leu Asn Leu Ala Leu Asn Ala Cys Gly Thr Leu Pro Asp Gly  
100 105 110

Phe Arg Cys Pro Thr  
115

(2) INFORMATION FOR SEQ ID NO:478:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 115 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..115  
(D) OTHER INFORMATION: / Ceres Seq. ID 2051327

574

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:478:

Met Ala Pro Lys Thr Ser Thr Thr Leu Ala Leu Phe Leu Val Thr Asn  
1 5 10 15

Ile Leu Phe Leu Asn Leu Ile Thr Leu Ser Cys Ala Asp Asn Thr Cys  
20 25 30

Pro Arg Asp Val Leu Lys Leu Ser Thr Cys Ser Asn Val Leu Asn Leu  
35 40 45

Ile Asn Leu Lys Leu Gly Ala Pro Ala Met Arg Pro Cys Cys Ser Ile  
50 55 60

Leu Phe Gly Leu Ile Asp Leu Asp Val Ala Val Cys Leu Cys Thr Ala  
65 70 75 80

Leu Cys Leu Ser Leu Leu Gly Ile Thr Ile Asp Thr Pro Ile His Leu  
85 90 95

Asn Leu Ala Leu Asn Ala Cys Gly Thr Leu Pro Asp Gly Phe Arg  
100 105 110

Cys Pro Thr  
115

(2) INFORMATION FOR SEQ ID NO:479:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 74 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..74  
(D) OTHER INFORMATION: / Ceres Seq. ID 2051328

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:479:

Met Leu Leu Phe Ala Phe Ala Pro Arg Ser Ser Ser Ala Phe Leu Ala  
1 5 10 15

Ser Pro Ser Thr Leu Leu Phe Thr Leu Thr Trp Leu Leu Thr Pro Val  
20 25 30

Glu Val Pro Phe Leu Met Asp Ser Val Ala Gln His Ser Tyr Asn Ile  
35 40 45

Phe Ile Phe Ser Ile Tyr Val Leu Glu Leu Val Leu Ser Asn Ile Gln  
50 55 60

Ile Ser Asn Leu Asn Ser Cys Ile Ser Lys  
65 70

(2) INFORMATION FOR SEQ ID NO:480:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 517 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..517  
(D) OTHER INFORMATION: / Ceres Seq. ID 2051633

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:480:

cataacccct ttggaagag ttcaacact tgcagagaaa aagaacaagg aagatcccg  
aaaatggcga cggcgatgtt acgttcagct ctttcccgag cagtgcactg cgcagctccg  
aagacatccg tgcctctcaa ggaaacttt tctcttccg ccggccatga cgtatcttat  
gtctatgttt tatccaaggg caatcatcac ggcaagacc ctctgccta tccgcatatg  
cacatccga acaaggagtt tcttggggt cggatggtc tggttaggt gaagacaaac  
aagagaccta gaacttcg tggatcaaat aagctcttt tggttattt gaaagctaa  
aatctttac cgtattgtt ctacagcttt gtaacgatt tgcacteca atctctttc  
tttgttggg aataaaagt taactttg cttgttc  
480



575

(A) LENGTH: 123 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..123

(C) OTHER INFORMATION: / Ceres Seq. ID 2051634

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:481:

His Asn Pro Leu Gly Lys Ser Leu Asn Thr Cys Arg Glu Lys Glu Gln

1 5 10 15

Gly Arg Ser Arg Lys Met Ala Thr Ala Ile Val Arg Ser Ala Leu Ser

20 25 30

Arg Ala Val Thr Arg Ala Ala Pro Lys Thr Ser Val Ala Pro Lys Arg

35 40 45

Asn Phe Ser Ser Ala Gly His Asp Asp Ala Tyr Glu Ala Ala Lys

50 55 60

Trp Glu Lys Ile Thr Tyr Leu Gly Ile Ala Ser Cys Thr Ala Leu Ala

65 70 75

Val Tyr Val Leu Ser Lys Gly His His His Gly Glu Asp Prc Pro Ala

80 85 90

Tyr Pro His Met His Ile Arg Asn Lys Glu Phe Pro Trp Gly Pro Asp

95 100 105

Gly Leu Phe Glu Val Lys His Asn Lys Glu His

110 115 120

(2) INFORMATION FOR SEQ ID NO:482:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 102 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..102

(C) OTHER INFORMATION: / Ceres Seq. ID 2051635

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:482:

Met Ala Thr Ala Ile Val Arg Ser Ala Leu Ser Arg Ala Val Thr Arg

1 5 10 15

Ala Ala Prc Lys Thr Ser Val Ala Pro Lys Arg Asn Phe Ser Ser

20 25 30

Ala Gly His Asp Asp Ala Tyr Glu Ala Ala Lys Trp Glu Lys Ile Thr

35 40 45

Tyr Leu Gly Ile Ala Ser Cys Thr Ala Leu Ala Val Tyr Val Leu Ser

50 55 60

Lys Gly His His His Gly Glu Asp Pro Pro Ala Tyr Pro His Met His

65 70 75

Ile Arg Asn Lys Glu Phe Pro Trp Gly Pro Asp Gly Leu Phe Glu Val

80 85 90

Lys His Asn Lys Lys Glu His

95 100

(2) INFORMATION FOR SEQ ID NO:483:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 410 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..410

576

(D) OTHER INFORMATION: / Ceres Seq. ID 2051906

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:483:

ttctctttt ttacatctta cccaaattc caaacacca tctgattttt gctctcttga

aatgtgtctt ttatgtcca attctctct tccaactaaa ccgaaccgga aaaccaggtt

cggagatcgg tgttactga tggccaaca gcaacgaacc cgaatttaca tacttcagc

atgtctctcc atgcgtcttt gctggcaca ccaattctat tctgattagt ttgattcag

aacggagcgg accaagaagt agaaacaga agaattgtgat gtcatttaga gttttcttg

tttttttcca gaactctca aggaattctt tgaattgaa gaagacctta aggatgggac

-ctttgggccc ttgtacata ttgttaaat aacatatgaa cttttttac

(2) INFORMATION FOR SEQ ID NO:484:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 75 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..75

(C) OTHER INFORMATION: / Ceres Seq. ID 2051907

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:484:

Ser Phe Phe Phe Thr Thr Gln Asn Ser Lys Thr Pro Ser Ile Phe

1 5 10 15

Cys Ser Leu Glu Met Cys Leu Phe Met Ser Asn Ser Ser Leu Pro Thr

20 25 30

Lys Pro Asn Arg Lys Thr Arg Phe Gly Asp Arg Cys Leu Met Ala

35 40 45

Lys Gln Gln Arg Thr Arg Leu Tyr Ile Leu Arg Arg Cys Val Ser Met

50 55 60

Leu Leu Cys Trp His Asp His Ser Ile Ser Asp

65 70 75

(2) INFORMATION FOR SEQ ID NO:485:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 55 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..55

(C) OTHER INFORMATION: / Ceres Seq. ID 2051908

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:485:

Met Cys Leu Phe Met Ser Asn Ser Ser Leu Pro Thr Lys Pro Asn Arg

1 5 10 15

Lys Thr Arg Phe Gly Asp Arg Cys Leu Met Ala Lys Gln Gln Arg

20 25 30

Thr Arg Leu Tyr Ile Leu Arg Arg Cys Val Ser Met Leu Leu Cys Trp

35 40 45

His Asp His Ser Ile Ser Asp

50 55

(2) INFORMATION FOR SEQ ID NO:486:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 51 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..51

(D) OTHER INFORMATION: / Ceres Seq. ID 2051909

577

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:486:

Met Ser Asn Ser Ser Leu Pro Thr Lys Pro Asn Arg Lys Thr Arg Phe  
 1 5 10 15  
 Gly Asp Arg Cys Leu Met Ala Lys Gln Gln Arg Thr Arg Leu Tyr  
 20 25 30  
 Ile Leu Arg Arg Cys Val Ser Met Leu Leu Cys Trp His Asp His Ser  
 35 40 45  
 Ile Ser Asp  
 50

(2) INFORMATION FOR SEQ ID NO:487:

(i) SEQUENCE CHARACTERISTICS:

(a) LENGTH: 719 base pairs  
 (b) TYPE: nucleic acid  
 (c) STRANDEDNESS: single  
 (d) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:

(a) NAME/KEY: -

(b) LOCATION: 1..719

(d) OTHER INFORMATION: / Ceres Seq. ID 2052403

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:487:

acaaacatt caaagacata caaataatt gaatttttt ttaactaga acaaaaggg  
 1 5 10 15  
 ggtgaatggg gttctcatg tggagttga ggttagttt cggctcgaa agttctggg  
 20 25 30  
 agccttggc gttcgatcca atctctccc caaagatttc ctaacgact acaaaacat  
 35 40 45  
 ccaagtcca cggcggaaga tatcggtccc tggctccatt cgcctatta ctatgaga  
 50 55 60  
 agaatctcca cggcggaaga tatcggtccc gaggatcgaa gcaggatatt tggagaacaa  
 65 70 75  
 aagcagtcg tacagatcca ttggcggaaga atgtctggg tactaataa ccttcaagg  
 80 85 90  
 aaccatccc gttacctta agatgtgtg tagctcttg aaagtgtg gtagattga  
 95 100 105  
 gaagaccgc catgagatcg atgatacca tgcatacaag gacttctg tcaagaact  
 110 115 120  
 taataagata gatggatgc tcttaagca aactagacc taactatga accttaaat  
 125 130 135  
 gaacccttt argaatatc agcttttatg attctctaa ttaagaagt gaataaagt  
 140 145 150  
 taagctttt acaaaagcca attagtcga acatttga aaaaatgaac cttgtgta  
 155 160 165

(2) INFORMATION FOR SEQ ID NO:488:

(i) SEQUENCE CHARACTERISTICS:

(a) LENGTH: 155 amino acids  
 (b) TYPE: amino acid  
 (c) STRANDEDNESS:  
 (d) TOPOLOGY: linear  
 (ix) FEATURE:

(a) NAME/KEY: peptide

(b) LOCATION: 1..155

(d) OTHER INFORMATION: / Ceres Seq. ID 2052404

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:488:

Met Gly Leu Ser Gly Val Leu His Val Glu Val Glu Val Lys Ser Pro  
 1 5 10 15  
 Ala Glu Lys Phe Trp Val Ala Leu Gly Asp Gly Ile Asn Leu Phe Pro  
 20 25 30  
 Lys Ala Phe Pro Asn Asp Tyr Lys Thr Ile Gln Val Leu Ala Gly Asp  
 35 40 45  
 Gly Asn Ala Pro Gly Ser Ile Arg Leu Ile Thr Tyr Gly Glu Gly Ser  
 50 55 60  
 Pro Leu Val Lys Ile Ser Ala Glu Arg Ile Glu Ala Val Asp Leu Glu  
 65 70 75  
 Asn Lys Ser Met Ser Tyr Ser Ile Ile Gly Gly Glu Met Leu Glu Tyr  
 80 85 90  
 Tyr Lys Thr Phe Lys Gly Thr Ile Thr Val Ile Pro Lys Asp Gly Gly  
 95 100 105  
 Ser Leu Leu Lys Trp Ser Gly Glu Phe Glu Lys Thr Ala His Glu Ile  
 110 115 120  
 Asp Asp Pro His Val Ile Lys Asp Phe Ala Val Lys Asn Phe Lys Glu  
 125 130 135

578

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:489:

Ile Asp Glu Tyr Leu Leu Lys Gln Thr Ser Ala  
 130 135 140 145 150 155

(2) INFORMATION FOR SEQ ID NO:489:

(a) LENGTH: 739 base pairs  
 (b) TYPE: nucleic acid  
 (c) STRANDEDNESS: single  
 (d) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:

(a) NAME/KEY: -

(b) LOCATION: 1..739

(d) OTHER INFORMATION: / Ceres Seq. ID 2053545

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:489:

atttagrac atgttgacc atcttttcg tataactac tatcttgat ctcttcgag  
 60  
 ttaagtcagt aactagaaa attcagaagc gctctcaatc tcaaaaatat ccatggcgc  
 120  
 gattacgaa ttctctacca aagatcacg atagtcgtt ctcgcctcg tcttctactg  
 180  
 ttctccaac ctctggatg gtcctcaagt cgtcagagt cgtcaaaagt acaactccc  
 240  
 gtatcact ctatcagaa tatcaaat cttagagct gatgcaarg taattcac tgaatctct  
 300  
 tcaagagga tatcaaat cttagagct gatgcaarg taattcac tgaatctct  
 360  
 cgttgatg aagcaccct gractgtac tggcttggc tgccttaca acgttcgag  
 420  
 atctctac tttaaggtt agctctcag agatccatg aagctctta cgtcgggaa  
 480  
 aacgtttc ttgggttgc tagctctat gatctgacc atctgttg gtcactct  
 540  
 gatctgct agctctacc gttctgggg ttaatgattc tctgttttc tcnagaata  
 600  
 tgaaccaa ctctgtaagc tgcacaaa acttgctaa tcttttagg tctgtcact  
 660  
 taaaagtt gaataaac atgctctat agaacgttg aaattcaca cctgtagacg  
 720  
 ttaacaaaga ttgaatta

(2) INFORMATION FOR SEQ ID NO:490:

(i) SEQUENCE CHARACTERISTICS:

(a) LENGTH: 146 amino acids  
 (b) TYPE: amino acid  
 (c) STRANDEDNESS:  
 (d) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

(a) NAME/KEY: peptide

(b) LOCATION: 1..146

(d) OTHER INFORMATION: / Ceres Seq. ID 2053546

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:490:

Met Ala Ala Ile Thr Glu Phe Leu Pro Lys Glu Tyr Gly Tyr Val Val  
 1 5 10 15  
 Leu Val Leu Val Phe Tyr Cys Phe Leu Asn Leu Trp Met Gly Ala Gln  
 20 25 30  
 Val Gly Arg Ala Arg Lys Arg Tyr Asn Val Pro Tyr Pro Thr Leu Tyr  
 35 40 45  
 Ala Ile Glu Ser Glu Asn Lys Asp Ala Lys Leu Phe Asn Cys Val Gln  
 50 55 60  
 Arg Gly His Gln Asn Ser Leu Glu Met Met Pro Met Tyr Phe Ile Leu  
 65 70 75  
 Met Ile Leu Gly Gly Met Lys His Pro Cys Ile Cys Thr Gly Leu Gly  
 80 85 90  
 Leu Leu Tyr Asn Val Ser Arg Phe Tyr Phe Lys Gly Tyr Ala Thr  
 95 100 105  
 Gly Asp Pro Met Lys Arg Leu Thr Ile Gly Lys Tyr Gly Phe Leu Gly  
 110 115 120  
 Leu Leu Gly Leu Met Ile Cys Thr Ile Ser Phe Gly Val Thr Leu Ile  
 125 130 135  
 Leu Ala  
 140 145

(2) INFORMATION FOR SEQ ID NO:491:

(i) SEQUENCE CHARACTERISTICS:

579

(A) LENGTH: 118 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ix) MOLECULE TYPE: peptide  
 (ix) FEATURE:

(A) NAME/KEY: peptide  
 (B) LOCATION: 1..118

(D) OTHER INFORMATION: / Ceres Seq. ID 2053547  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:491:

Met Gly Ala Gln Val Gly Arg Ala Arg Lys Arg Tyr Asn Val Pro Tyr  
 1 5 10 15

Pro Thr Leu Tyr Ala Ile Glu Ser Glu Asn Lys Asp Ala Lys Leu Phe  
 20 25 30

Asn Cys Val Gln Arg Gly His Gln Asn Ser Leu Glu Met Met Pro Met  
 35 40 45

Tyr Phe Ile Leu Met Ile Leu Gly Gly Met Lys His Pro Cys Ile Cys  
 50 55 60

Thr Gly Leu Gly Leu Leu Tyr Asn Val Ser Arg Phe Phe Tyr Phe Lys  
 65 70 75 80

Gly Tyr Ala Thr Gly Asp Pro Met Lys Arg Leu Thr Ile Gly Lys Tyr  
 85 90 95

Gly Phe Leu Gly Leu Leu Gly Leu Met Ile Cys Thr Ile Ser Phe Gly  
 100 105 110

Val Thr Leu Ile Leu Ala  
 115

(2) INFORMATION FOR SEQ ID NO:492:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 74 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ix) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..74

(D) OTHER INFORMATION: / Ceres Seq. ID 2053548  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:493:

Met Met Pro Met Tyr Phe Ile Leu Met Ile Leu Gly Gly Met Lys His  
 1 5 10

Pro Cys Ile Cys Thr Gly Leu Gly Leu Tyr Asn Val Ser Arg Phe  
 20 25 30

Phe Tyr Phe Lys Gly Tyr Ala Thr Gly Asp Pro Met Lys Arg Leu Thr  
 35 40 45

Ile Gly Lys Tyr Gly Phe Leu Gly Leu Leu Gly Leu Met Ile Cys Thr  
 50 55 60

Ile Ser Phe Gly Val Thr Leu Ile Leu Ala  
 65

(2) INFORMATION FOR SEQ ID NO:493:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 681 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..681

(D) OTHER INFORMATION: / Ceres Seq. ID 2053984  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:493:

ccatccaaataaaccaaac aaaaattat attcaagaga aaaaaggaataa aatgaatttc  
 60  
 attccgagtc aggttaagaa actctcaagc tcaacaccag agggagccaga cccaaacaaag  
 120

580

ccagtcgaag gaaccgaac agctacaaga ccagctacca acgcagagct catggcaagt  
 180  
 gcaaggttg tagctgaagc tgctcaagcc gcagctcgta acgaatcaga caaatcgac  
 240  
 aaggttaag tcgccggagc ctctgcgat attctagag ctgcgaagaa atacgttaag  
 300  
 ttctgataaa agagttagac tggctagtac ctgcacaagg ctgagaagta tctcaagac  
 360  
 tagagtggt cacactcac cggctcgggt ggtctctctc ctccgacgag tcaggtcgag  
 420  
 ccaggaagtc agctcgagcc ggcggtcaag aagaagatg aagatctgg tgggtgctt  
 480  
 ggaggttatg ccaatagcc tcaaggtttc ttgaagtat ttgatctta attgtgttc  
 540  
 cgttatggt tatgaggaat gacagtgag tgaataaact agtatggtt gtaactagt tagtgttct  
 600  
 tccatcttcg tctgattcat g

(2) INFORMATION FOR SEQ ID NO:494:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 172 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ix) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..172

(D) OTHER INFORMATION: / Ceres Seq. ID 2053885

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:494:

Pro Ser Lys Lys Asn Lys Thr Lys Asn Tyr Ile Gln Glu Lys Lys Gly  
 1 5 10 15

Lys Met Asn Phe Ile Ser Asp Gln Val Lys Lys Lys Leu Ser Ser Thr  
 20 25 30

Pro Glu Glu Pro Asp His Asn Lys Pro Val Glu Gly Thr Glu Thr Ala  
 35 40 45

Thr Arg Pro Ala Thr Asn Ala Glu Leu Met Ala Ser Ala Lys Val Val  
 50 55 60

Ala Glu Ala Ala Gln Ala Ala Arg Asn Glu Ser Asp Lys Leu Asp  
 65 70 75

Lys Gly Lys Val Ala Gly Ala Ser Ala Asp Ile Leu Asp Ala Ala Glu  
 80 85 90 95

Lys Tyr Gly Lys Phe Asp Glu Lys Ser Thr Gly Gln Tyr Leu Asp  
 100 105 110

Lys Ala Glu Lys Tyr Leu Asn Asp Tyr Glu Ser Ser His Ser Thr Gly  
 115 120 125

Ala Gly Gly Pro Pro Pro Thr Ser Gln Ala Glu Pro Ala Ser Gln  
 130 135 140

Pro Glu-Pro Ala Ala Lys Lys Asp Asp Glu Ser Gly Gly Gly Leu  
 145 150 155

Gly Gly Tyr Ala Lys Met Ala Gln Gly Phe Leu Lys  
 160 165 170

(2) INFORMATION FOR SEQ ID NO:495:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 155 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ix) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..155

(D) OTHER INFORMATION: / Ceres Seq. ID 2053886

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:495:

Met Asn Phe Ile Ser Asp Gln Val Lys Lys Leu Ser Ser Ser Thr Pro  
 1 5 10 15

Glu Glu Pro Asp His Asn Lys Pro Val Glu Gly Thr Glu Thr Ala Thr  
 20 25 30

Arg Pro Ala Thr Asn Ala Glu Leu Met Ala Ser Ala Lys Val Val Ala  
 35 40 45

581

Glu Ala Ala Gln Ala Ala Arg Asn Glu Ser Asp Lys Leu Asp Lys  
50 55 60  
Gly Lys Val Ala Gly Ala Ser Ala Asp Ile Leu Asp Ala Ala Glu Lys  
65 70 75 80  
Tyr Gly Lys Phe Asp Glu Lys Ser Ser Thr Gly Gln Tyr Leu Asp Lys  
85 90 95  
Ala Glu Lys Tyr Leu Asn Asp Tyr Glu Ser Ser His Ser Thr Gly Ala  
100 105 110  
Gly Gly Pro Pro Pro Thr Ser Gln Ala Glu Pro Ala Ser Gln Pro  
115 120 125  
Glu Pro Ala Ala Lys Lys Asp Asp Glu Glu Ser Gly Gly Gly Leu Gly  
130 135 140  
Gly Tyr Ala Lys Met Ala Gln Gly Phe Leu Lys  
145 150 155

5

10

15

20

25

30

35

40

45

50

55

60

tctctcdaa atccataaaa agagagagag ataaataaag aaaaactaa agaagctaga  
120  
agatggaaa aagtaaatgat catgacaag catgacacgg cggctccggt ggtgtgcca  
180  
cggagaagtg gagagagact agctcggaaa tccgaactgc cggacacatg cccgggtg  
240  
ctcgggtgg ccttggtgtt ggaacacttg ttgtcactgt taagactct gagactaag  
300  
agtcgggttc attctctac tccattcca cagcttttag gtacttggtg cagcgaag  
360  
gaatgtgc aggtactact cttctacag cagccatgc agcgatgct cgtctctt  
420  
cgaatgac tegtgttg accctttt gtccagaca gttctgacc taccgtgtc  
480  
ttctgctgg agctgtaca gctgaggttc taccatggg gactcagcca

582

ttactggaag caatgatgt agctctatg gcgtttctg tcatagacc actgtctctg  
540  
ttataacac attcttggtg gttgtttct acatcttct cttctaatc tctcttata  
600  
agctcttac tgccttgat cctctccca ttgtgacc cgcagaagt ctcgaagtcg  
660  
ctctctcgg aagttagat cctcattgc tctcaagta cctatgttt gcatgttcc  
720  
aaatgtgtg tgttttaac ttgtgtcaag agaagaatg cttatgtgt tgccttctt  
780  
tcattgctt tctctt

(2) INFORMATION FOR SEQ ID NO:498:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 204 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..204

(D) OTHER INFORMATION: / Ceres Seq. ID 2053909

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:498:

Met Glu Lys Ser Asn Asp His Asp Lys Ala Ser His Gly Gly Ser Gly  
1 5 10 15

Gly Gly Ala Thr Glu Lys Trp Glu Glu Thr Ser Leu Gly Ile Arg Thr  
20 25 30

Ala Glu Thr Met Leu Arg Leu Ala Pro Val Gly Leu Cys Val Ala Ala  
35 40 45

Leu Val Val Met Leu Lys Asp Ser Glu Thr Asn Glu Phe Gly Ser Ile  
50 55 60

Ser Tyr Ser Asn Leu Thr Ala Phe Arg Tyr Leu Val His Ala Asn Gly  
65 70 75 80

Ile Cys Ala Gly Tyr Ser Leu Leu Ser Ala Ala Ile Ala Ala Met Pro  
85 90 95

Arg Ser Ser Ser Thr Met Pro Arg Val Trp Thr Phe Phe Cys Leu Asp  
100 105 110

Gln Leu Leu Thr Tyr Leu Val Leu Ala Ala Gly Ala Val Ser Ala Glu  
115 120 125

Val Leu Tyr Leu Ala Tyr Asn Gly Asp Ser Ala Ile Thr Trp Ser Asp  
130 135 140

Ala Cys Ser Ser Tyr Gly Gly Phe Cys His Arg Ala Thr Ala Ser Val  
145 150 155

Ile Ile Thr Phe Phe Val Val Cys Phe Tyr Ile Val Leu Ser Leu Ile  
160 165 170 175

Ser Ser Tyr Lys Leu Phe Thr Arg Phe Asp Pro Ser Ile Val Asp  
180 185 190

Ser Ala Lys Asn Leu Glu Val Ala Val Phe Gly Ser  
195 200

(2) INFORMATION FOR SEQ ID NO:499:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 169 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..169

(D) OTHER INFORMATION: / Ceres Seq. ID 2053910

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:499:

Met Leu Arg Leu Ala Pro Val Gly Leu Cys Val Ala Ala Leu Val Val  
1 5 10 15

Met Leu Lys Asp Ser Glu Thr Asn Glu Phe Gly Ser Ile Ser Tyr Ser  
20 25 30

Asn Leu Thr Ala Phe Arg Tyr Leu Val His Ala Asn Gly Ile Cys Ala  
35 40 45

583

Gly Tyr Ser Leu Leu Ser Ala Ala Ile Ala Met Pro Arg Ser Ser  
50 55 60  
Ser Thr Met Pro Arg Val Trp Thr Phe Phe Cys Leu Asp Gln Leu Leu  
65 70 75 80  
5 Thr Tyr Leu Val Leu Ala Ala Gly Ala Val Ser Ala Glu Val Leu Tyr  
85 90 95  
Leu Ala Tyr Asn Gly Asp Ser Ala Ile Thr Trp Ser Asp Ala Cys Ser  
100 105 110  
10 Ser Tyr Gly Gly Phe Cys His Arg Ala Thr Ala Ser Val Ile Ile Thr  
115 120 125  
Phe Phe Val Val Cys Phe Tyr Ile Val Leu Ser Leu Ile Ser Ser Tyr  
130 135 140  
Lys Leu Phe Thr Arg Phe Asp Pro Pro Ser Ile Val Asp Ser Ala Lys  
145 150 155 160  
15 Asn Leu Glu Val Ala Val Phe Gly Ser  
165

(2) INFORMATION FOR SEQ ID NO:500:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 153 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) MOLECULE TYPE: peptide  
(ix) FEATURE:

- (A) NAME/KEY: peptide  
(B) LOCATION: 1..153  
(D) OTHER INFORMATION: / Ceres Seq. ID 2053911

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:500:

Met Leu Lys Asp Ser Glu Thr Asn Glu Phe Gly Ser Ile Ser Tyr Ser  
1 5 10 15  
Asn Leu Thr Ala Phe Arg Tyr Leu Val His Ala Asn Gly Ile Cys Ala  
20 25 30  
Gly Tyr Ser Leu Leu Ser Ala Ala Ile Ala Ala Met Pro Arg Ser Ser  
35 40 45  
Ser Thr Met Pro Arg Val Trp Thr Phe Phe Cys Leu Asp Gln Leu Leu  
50 55 60  
Thr Tyr Leu Val Leu Ala Ala Gly Ala Val Ser Ala Glu Val Leu Tyr  
65 70 75 80  
Leu Ala Tyr Asn Gly Asp Ser Ala Ile Thr Trp Ser Asp Ala Cys Ser  
85 90 95  
40 Ser Tyr Gly Gly Phe Cys His Arg Ala Thr Ala Ser Val Ile Ile Thr  
100 105 110  
Phe Phe Val Val Cys Phe Tyr Ile Val Leu Ser Leu Ile Ser Ser Tyr  
115 120 125  
45 Lys Leu Phe Thr Arg Phe Asp Pro Pro Ser Ile Val Asp Ser Ala Lys  
130 135 140  
Asn Leu Glu Val Ala Val Phe Gly Ser  
145 150

(2) INFORMATION FOR SEQ ID NO:501:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 652 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ix) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

- (A) NAME/KEY: -  
(B) LOCATION: 1..652  
(D) OTHER INFORMATION: / Ceres Seq. ID 2056123

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:501:

atttttttcttctctctcccttaagc aaactaaaa caagctatgg ctggtatgct  
60  
tcccgagatt gagtgtgcaa ggaggggcg cttccacggt ggtgtctctc cgattgaatc  
120

584

ctcgaaacaca gttctgtgg oggtgcggc gggaacagtc tggacaagcgc gaccatcgtt  
ctcttttac attaccaatc atgagagcca ccagggccat gttctctctt cggagagaag  
tgttaggaat aatcttatg gagaagacaa cgtatgagaa cttgacggag cagccaaga  
ggcaagcag aggttaaca agcgtgtgg agtccacca cgtacaagc aaatgtgtaa  
agacaaggg aataattgg agcaagaaa ggttaacct ctggggact taccagcca  
gggtcggg ttaagaaga gccaggaag gttagtgaa tggttcaagc ggcaggttag  
ggacaacaa gattgtgcta tatctctaga ccggttcaag aaggttgaga cctgggaca  
ccacaagct gccataagt ttactccat atgcttatg cttgtgctag acacaatgt  
tatggcca tatgtaga ctgatattg gaattaatg ttatatattt ga  
20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195

(2) INFORMATION FOR SEQ ID NO:502:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 196 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) MOLECULE TYPE: peptide  
(ix) FEATURE:

- (A) NAME/KEY: peptide  
(B) LOCATION: 1..196  
(D) OTHER INFORMATION: / Ceres Seq. ID 2056124

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:502:

Met Ala Gly Met Leu Pro Gly Val Glu Cys Ala Arg Arg Arg Arg Phe  
1 5 10 15  
His Gly Gly Ala Pro Pro Ile Glu Ser Ser Asn Thr Ala Ser Val Ala  
20 25 30 35 40 45  
Ala Ala Ala Gly His Val Trp Thr Arg Arg Pro Ser Phe Ser Leu Tyr  
Thr Thr Asn His Glu Ser His Gln Ala His Val Ser Phe Ser Glu Arg  
50 55 60 65 70 75 80 85 90 95  
Ser Val Arg Asn Lys Ser Tyr Gly Glu Asp Asn Asp Glu Lys Leu Asp  
Gly Ala Ala Lys Glu Ala Lys Gln Arg Leu Asn Lys Arg Leu Arg Ile  
Pro Pro Arg Thr Arg Gln Asn Gly Lys Asp Lys Gly Asn Lys Leu Glu  
Gln Gly Lys Gly Lys Pro Leu Gly Asp Leu Pro Thr Glu Val Val Gly  
Leu Lys Lys Ser Arg Gly Arg Leu Met Glu Trp Phe Lys Arg Arg Val  
Arg Glu Gln Gln Asp Cys Ala Ile Cys Leu Asp Arg Phe Lys Lys Gly  
Glu Thr Leu Val His Leu Pro Cys Ala His Lys Phe His Ser Ile Cys  
Leu Leu Pro Trp Leu Asp Thr Asn Val Tyr Cys Pro Tyr Cys Arg Thr  
Asp Ile Trp Asn  
195

(2) INFORMATION FOR SEQ ID NO:503:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 193 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ix) MOLECULE TYPE: peptide  
(ix) FEATURE:

- (A) NAME/KEY: peptide  
(B) LOCATION: 1..193  
(D) OTHER INFORMATION: / Ceres Seq. ID 2056125

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:503:

Met Leu Pro Gly Val Glu Cys Ala Arg Arg Arg Arg Phe His Gly Gly  
1 5 10 15  
Ala Pro Pro Ile Glu Ser Ser Asn Thr Ala Ser Val Ala Ala Ala  
190

586

accgaacagc aagttatgaa ccgaaacgg tttttggaag aagagcgga agcgagggg 1680  
 ggaaaggtt tagtgctcg tgggttaaca tggaaatca gtttccaag gtttcgagcg 1740  
 ggtcatgta aggtttcat ggaatcagc gatgtggac gcaactcga tcttcggtt 1800  
 atggcgctg accaagaatt gaccggaaa ttgctcgaga ttttcctaat agaaagagc 1920  
 tggagattt tgaatcgtt tgttgcagg gatcaaatg gtttcaatc acgtattgga 1980  
 gagcaacctt taagtatt catgaagca actaaacggc taacataca gatgatatt 2040  
 gttggaca acgtgagaaa catttttagt tctctcaa taatatttt tctatttt 2100  
 tcaaaaat taattttt ttgttttt ttccggtg gatcactac cctacatc 2160  
 atataatt cagatatat atataatat atataatat atataatc atataaata 2220  
 aaactttt ttgtcatt tttaagttt aaacagtc aataatc atacttat accattcg 2280  
 atgaagaac taataaac ctgttcaat gttatgata tctcaaat tcttgctca 2340  
 ggaggggg caatatagg gtaggaaa ccaaaaaa aatgaatga gattcaag 2400  
 caataact ggttgctc agagtgcac tctatgtg taagaagag ggggaaga 2460  
 gccgttct gctctctc taataagtt tactgtatt acattgcc taacattt 2520  
 tattattac aataataga gattatgt aagatgaat aataatga cgggtgta 2580  
 tagcccaag atgatatg ttgtctat caataaac atataatc atataatga aagaataa

(1) INFORMATION FOR SEQ ID NO:504:  
 (A) LENGTH: 2619 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..2619

(D) OTHER INFORMATION: / Ceres Seq. ID 2056245

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:504:  
 Met Glu Gln Glu Lys Ser Leu Asp Pro Gln Leu Thr His Ala Cys Ala  
 Gly Ser Met Val Gln Ile Pro Ser Leu Asn Ser Thr Val Phe Tyr Phe  
 Ala Gln Gly His Thr Glu His Ala His Ala Pro Pro Asp Phe His Ala  
 Pro Arg Val Pro Pro Leu Ile Leu Cys Arg Val Val Ser Val Lys Phe  
 Leu Ala Asp Ala Glu Thr Asp Glu Val Phe Ala Lys Ile Thr Leu Leu  
 Pro Leu Pro Gly Asn Asp Leu Asp Leu Glu Asn Asp Ala Val Leu Gly  
 Leu Thr Pro Pro Ser Ser Asp Gly Asn Gly Asn Gly Lys Lys Pro  
 Ala Ser Phe Ala Lys Thr Leu Thr Gln Ser Asp Ala Asn Asn Gly Gly  
 Gly Phe Ser Val Pro Arg Tyr Cys Ala Glu Thr Ile Phe Pro Arg Leu  
 Asp Tyr Ser Ala Glu Pro Pro Val Gln Thr Val Ile Ala Lys Asp Ile  
 His Gly Glu Thr Thr Lys Phe Arg His Ile Tyr Arg Gly Thr Pro Arg  
 Arg His Leu Leu Thr Thr Gly Thr Ser Thr Phe Val Asn Gln Lys Lys  
 Leu Ile Ala Gly Asp Ser Ile Val Phe Leu Arg Ser Glu Ser Gly Asp  
 Leu Cys Val Gly Ile Arg Arg Ala Lys Arg Gly Gly Leu Gly Ser Asn  
 Ala Gly Ser Asp Asn Pro Tyr Pro Gly Phe Ser Gly Phe Leu Arg Asp  
 Asp Glu Ser Thr Thr Thr Ser Lys Leu Met Met Met Lys Arg Asn  
 Gly Asn Asn Asp Gly Asn Ala Ala Ala Thr Gly Arg Val Arg Val Glu

585

Gly His Val Trp Thr Arg Arg Pro Ser Phe Ser Leu Tyr Thr Asn 30  
 35 40 45  
 His Glu Ser His Gln Ala His Val Ser Phe Ser Glu Arg Ser Val Arg 45  
 50 55 60  
 Asn Lys Ser Tyr Gly Glu Asp Asn Asp Glu Lys Leu Asp Gly Ala Ala 60  
 65 70 75  
 Lys Glu Ala Lys Gln Arg Leu Asn Lys Arg Leu Arg Ile Pro Pro Arg 80  
 85 90 95  
 Thr Arg Gln Asn Gly Lys Asp Lys Gly Asn Lys Leu Glu Gln Gly Lys 100  
 105 110 115  
 Gly Lys Pro Leu Gly Asp Leu Pro Thr Glu Val Val Gly Leu Lys Lys 120  
 125 130  
 Ser Arg Gly Arg Leu Met Glu Trp Phe Lys Arg Arg Val Arg Glu Gln 135  
 140 145  
 Gln Asp Cys Ala Ile Cys Leu Asp Arg Phe Lys Lys Gly Glu Thr Leu 150  
 155 160  
 Val His Leu Pro Cys Ala His Lys Phe His Ser Ile Cys Leu Leu Pro 165  
 170 175  
 Trp Leu Asp Thr Asn Val Tyr Cys Pro Tyr Cys Arg Thr Asp Ile Trp 180  
 185 190  
 Asn

(2) INFORMATION FOR SEQ ID NO:504:  
 (A) LENGTH: 2619 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..2619

(D) OTHER INFORMATION: / Ceres Seq. ID 2056245

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:504:  
 atggagaagc agaaagctt ggtacacaa ctatggatg tttgcagg atcaatggt 60  
 caaatcctt cactgaattc acggttttt tacttcctc aaggcacac agagacgct 120  
 caacggctc ctgatttca cggcgccgc gttcacctc ttactctct tccgtctgt 180  
 tccgtgaat tctctcgca cgttgaaca gagaagttt ttgttaaat tccgttttg 240  
 tcccttcg gaaacgatt ggtattaga aagcagcgg ttgttggtt acctctcct 300  
 tctctcag ctaacgttaa cgttaagag aacacggcg ctttcgctaa acgttaagc 360  
 cagctcagc ctataacgg cgggtgttct tccgttccac gttatggc cagagtaagt 420  
 tttccggcg ttgattact gggagacca cgggttcaa ccgtgatgc taagacatc 480  
 cagcggaga cttggaatt cggatattc tacaaggaa caactcgcg taactctca 540  
 accacgggt ggaacactt cgttaacag aagaactaa tgcgcgaga ctcaacgic 600  
 tttctcgtt ctgaacccg tgaactcgc gttgaacct gacgcgttaa accggcggt 660  
 ctcggatct accgagctt tacccttgt tctcgttgt tctcgttgt ctcctgtgc 720  
 gacagtcac caacacacac atgaagata atgatgatg aacgcagcg aaacaacag 780  
 ggaaacgcg cgcctacag gaggtttaga tagaagcag tagcgaagc gttggcggt 840  
 ccagcgttg gacaagcgt ttgattgtt tattatcac gcgtatgac accgagttt 900  
 tgcgaaay cagctgatgt tagatcaga atggagata gatgggtag tggatgcgt 960  
 tttaaatg cgtttgaac agagattct ttgaagata gttgtttt ggtatgctc 1020  
 tccgcgttc adgcctgta tccaattgt tggcattgt caccatggc tctcttcag 1080  
 ganttgat cgaacatgc taataatc taatctcat tctcttcag gaaagaatt 1140  
 aggtattcc agcatatga gttctatc accgttact aattccgat tttctcccg 1200  
 ggtattcca accatggcg ttgcagctc atgtgtatc tgtcaacga caacaata 1260  
 gctctcagc gaatacagc agcagagca gctcaacac tttcggatc acctctccg 1320  
 tctctgtct ctatctcaa tcttagagt taacacgta acaaagtt acctctccg 1380  
 gcgatgttc tatcgagtt caaccggcg catcattc accaggctag ggaatggcg 1440  
 aatgataa acattcggc ttttaact atgggaatc cgtctagt tctgataag 1500  
 aagaaagt ttgttcgt tagactcat cagttcgtt ttgttcgta accgattta 1560  
 1620

		587	
5	Ala Val	260	265
	Glu Ala Val	275	280
	Tyr Tyr Pro Arg	295	300
	Ala Val	310	315
	Phe Lys Met Ala Phe Glu Thr Glu Asp Ser Ser Arg Ile Ser Trp Phe	325	330
10	Met Gly Thr Val Ser Ala Val Gln Val Ala Asp Pro Ile Arg Trp Pro	340	345
	Asn Ser Pro Trp Arg Leu Leu Gln Val Ala Trp Asp Glu Pro Asp Leu	355	360
	Leu Gln Asn Val Lys Arg Val Ser Pro Trp Leu Val Glu Leu Val Ser	370	375
	Asn Met Pro Thr Ile His Leu Ser Pro Phe Ser Pro Arg Lys Lys Ile	385	390
	Arg Ile Pro Gln Pro Phe Glu Phe Pro Phe His Gly Thr Lys Phe Pro	405	410
15	Ile Phe Ser Pro Gly Phe Ala Asn Asn Gly Gly Glu Ser Met Cys	420	425
	Tyr Leu Ser Asn Asp Asn Asn Ala Pro Ala Gly Ile Gln Gly Ala	435	440
	Arg Gln Ala Gln Gln Leu Phe Gly Ser Pro Ser Pro Ser Leu Leu Ser	450	455
	Asp Leu Asn Leu Ser Ser Tyr Thr Gly Asn Asn Lys Leu His Ser Pro	465	470
	Ala Met Phe Leu Ser Ser Phe Asn Pro Arg His His His Tyr Gln Ala	485	490
20	Arg Asp Ser Glu Asn Ser Asn Ile Ser Cys Ser Leu Thr Met Gly	500	505
	Asn Pro Ala Met Val Gln Asp Lys Lys Ser Val Gly Ser Val Lys	515	520
	Thr His Gln Phe Val Leu Phe Gly Gln Pro Ile Leu Thr Glu Gln Gln	530	535
	Val Met Asn Arg Lys Arg Phe Leu Glu Glu Glu Ala Glu Glu Glu	545	550
	Glu Lys Gly Leu Val Ala Arg Gly Leu Thr Trp Asn Tyr Ser Leu Gln	565	570
25	Gly Leu Glu Thr Gly His Cys Lys Val Phe Met Glu Ser Glu Asp Val	580	585
	Gly Arg Thr Leu Asp Leu Ser Val Ile Gly Ser Tyr Gln Glu Leu Tyr	595	600
	Arg Lys Leu Ala Glu Met Phe His Ile Glu Glu Arg Ser Asp Leu Leu	610	615
	Thr His Val Val Tyr Arg Asp Ala Asn Gly Val Ile Lys Arg Ile Gly	625	630
	Asp Glu Pro Phe Ser Asp Phe Met Lys Ala Thr Lys Arg Leu Thr Ile	645	650
30	Lys Met Asp Ile Gly Gly Asp Asn Val Arg Asn Ile Phe Ser Phe Leu	660	665
	Gln Ile Ile Phe Phe Phe Phe Lys Lys Leu Ile Phe Phe Leu	675	680
	Phe Phe Ser Asp Gly Tyr Tyr Thr His Ile His Ile Tyr Ser Ser	690	695
	Asp Tyr Ile Tyr Ile Tyr Ile Tyr Ile Ser Asn Ile Lys Lys	705	710
	Lys Leu Ile Leu Phe Ile Leu Leu Ser Leu Lys Thr Val Asn Ser Phe	725	730
35	Ile Ser Ser Met Met Lys Lys Leu Ile Gln Leu Ser Phe Thr Val Met	740	745
	Ile Ile Phe Thr Ile Leu Val Leu Gly Val Ala Asn Glu Gly Leu	755	760
	Gly Lys Pro Lys Lys Gln Cys Asn Glu Ile Leu Lys Gln Ser Asn Cys	770	775
	Val Ala Ala Glu Cys Asp Ser Met Cys Val Lys Lys Arg Gly Lys Gly	785	790
	Ala Gly Tyr Cys Ser Pro Ser Lys Lys Cys Tyr Tyr His Cys	805	810
		588	
10	Gly Lys Pro Lys Lys Gln Cys Asn Glu Ile Leu Lys Gln Ser Asn Cys	755	760
	Val Ala Ala Glu Cys Asp Ser Met Cys Val Lys Lys Arg Gly Lys Gly	770	775
	Ala Gly Tyr Cys Ser Pro Ser Lys Lys Cys Tyr Tyr His Cys	785	790
	Pro	805	810
		589	
		590	
		591	
		592	
		593	
		594	
		595	
		596	
		597	
		598	
		599	
		600	
		601	
		602	
		603	
		604	
		605	
		606	
		607	
		608	
		609	
		610	
		611	
		612	
		613	
		614	
		615	
		616	
		617	
		618	
		619	
		620	
		621	
		622	
		623	
		624	
		625	
		626	
		627	
		628	
		629	
		630	
		631	
		632	
		633	
		634	
		635	
		636	
		637	
		638	
		639	
		640	
		641	
		642	
		643	
		644	
		645	
		646	
		647	
		648	
		649	
		650	
		651	
		652	
		653	
		654	
		655	
		656	
		657	
		658	
		659	
		660	
		661	
		662	
		663	
		664	
		665	
		666	
		667	
		668	
		669	
		670	
		671	
		672	
		673	
		674	
		675	
		676	
		677	
		678	
		679	
		680	
		681	
		682	
		683	
		684	
		685	
		686	
		687	
		688	
		689	
		690	
		691	
		692	
		693	
		694	
		695	
		696	
		697	
		698	
		699	
		700	
		701	
		702	
		703	
		704	
		705	
		706	
		707	
		708	
		709	
		710	
		711	
		712	
		713	
		714	
		715	
		716	
		717	
		718	
		719	
		720	
		721	
		722	
		723	
		724	
		725	
		726	
		727	
		728	
		729	
		730	
		731	
		732	
		733	
		734	
		735	
		736	
		737	
		738	
		739	
		740	
		741	
		742	
		743	
		744	
		745	
		746	
		747	
		748	
		749	
		750	
		751	
		752	
		753	
		754	
		755	
		756	
		757	
		758	
		759	
		760	
		761	
		762	
		763	
		764	
		765	
		766	
		767	
		768	
		769	
		770	
		771	
		772	
		773	
		774	
		775	
		776	
		777	
		778	
		779	
		780	





5	591	Lys Leu Ile Phe Phe Leu Phe Phe Ser Asp Gly Gly Tyr Tyr Thr His	420 425 430 440 445
		Ile His Ile Tyr Ser Ser Asp Tyr Ile Tyr Ile Tyr Ile Tyr	450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560
10	592	Leu Ser Phe Thr Val Met Ile Ile Phe Thr Ile Leu Val Leu Gly Val	500 505 510 515 520 525 530 535 540 545 550 555 560
		Val Ala Asn Glu Gly Leu Gly Lys Pro Lys Lys Lys Lys Lys Lys Lys	565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 900 905 910 915 920 925 930 935 940 945 950 955 960 965 970 975 980 985 990 995
15	593	Leu Lys Gln Ser Asn Cys Val Ala Ala Glu Cys Asp Ser Met Cys Val	960 965 970 975 980 985 990 995
		Lys Lys Arg Gly Lys Gly Ala Gly Tyr Cys Ser Pro Ser Lys Lys Cys	1000 1005 1010 1015 1020 1025 1030 1035 1040 1045 1050 1055 1060 1065 1070 1075 1080 1085 1090 1095 1100 1105 1110 1115 1120 1125 1130 1135 1140 1145 1150 1155 1160 1165 1170 1175 1180 1185 1190 1195 1200 1205 1210 1215 1220 1225 1230 1235 1240 1245 1250 1255 1260 1265 1270 1275 1280 1285 1290 1295 1300 1305 1310 1315 1320 1325 1330 1335 1340 1345 1350 1355 1360 1365 1370 1375 1380 1385 1390 1395 1400 1405 1410 1415 1420 1425 1430 1435 1440 1445 1450 1455 1460 1465 1470 1475 1480 1485 1490 1495 1500 1505 1510 1515 1520 1525 1530 1535 1540 1545 1550 1555 1560 1565 1570 1575 1580 1585 1590 1595 1600 1605 1610 1615 1620 1625 1630 1635 1640 1645 1650 1655 1660 1665 1670 1675 1680 1685 1690 1695 1700 1705 1710 1715 1720 1725 1730 1735 1740 1745 1750 1755 1760 1765 1770 1775 1780 1785 1790 1795 1800 1805 1810 1815 1820 1825 1830 1835 1840 1845 1850 1855 1860 1865 1870 1875 1880 1885 1890 1895 1900 1905 1910 1915 1920 1925 1930 1935 1940 1945 1950 1955 1960 1965 1970 1975 1980 1985 1990 1995
20	594	Tyr Cys Tyr Tyr His Cys Pro	2000 2005 2010 2015 2020 2025 2030 2035 2040 2045 2050 2055 2060 2065 2070 2075 2080 2085 2090 2095 2100 2105 2110 2115 2120 2125 2130 2135 2140 2145 2150 2155 2160 2165 2170 2175 2180 2185 2190 2195 2200 2205 2210 2215 2220 2225 2230 2235 2240 2245 2250 2255 2260 2265 2270 2275 2280 2285 2290 2295 2300 2305 2310 2315 2320 2325 2330 2335 2340 2345 2350 2355 2360 2365 2370 2375 2380 2385 2390 2395 2400 2405 2410 2415 2420 2425 2430 2435 2440 2445 2450 2455 2460 2465 2470 2475 2480 2485 2490 2495 2500 2505 2510 2515 2520 2525 2530 2535 2540 2545 2550 2555 2560 2565 2570 2575 2580 2585 2590 2595 2600 2605 2610 2615 2620 2625 2630 2635 2640 2645 2650 2655 2660 2665 2670 2675 2680 2685 2690 2695 2700 2705 2710 2715 2720 2725 2730 2735 2740 2745 2750 2755 2760 2765 2770 2775 2780 2785 2790 2795 2800 2805 2810 2815 2820 2825 2830 2835 2840 2845 2850 2855 2860 2865 2870 2875 2880 2885 2890 2895 2900 2905 2910 2915 2920 2925 2930 2935 2940 2945 2950 2955 2960 2965 2970 2975 2980 2985 2990 2995
		(2) INFORMATION FOR SEQ ID NO:508:	
25	595	(i) SEQUENCE CHARACTERISTICS:	
		(A) LENGTH: 776 base pairs	
30	596	(B) TYPE: nucleic acid	
		(C) STRANDEDNESS: single	
35	597	(D) TOPOLOGY: linear	
		(ii) MOLECULE TYPE: DNA (genomic)	
40	598	(ix) FEATURE:	
		(A) NAME/KEY: -	
45	599	(B) LOCATION: 1..776	
		(D) OTHER INFORMATION: / Ceres Seq. ID 1941142	
50	600	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:509:	
		agtcgctcc ttcctggcc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc gcgcgcgcgc	600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 900 905 910 915 920 925 930 935 940 945 950 955 960 965 970 975 980 985 990 995
55	601	(2) INFORMATION FOR SEQ ID NO:509:	
		(i) SEQUENCE CHARACTERISTICS:	
60	602	(A) LENGTH: 161 amino acids	
		(B) TYPE: amino acid	
65	603	(C) STRANDEDNESS:	
		(D) TOPOLOGY: linear	
70	604	(ii) MOLECULE TYPE: peptide	
		(ix) FEATURE:	
75	605	(A) NAME/KEY: peptide	
		(B) LOCATION: 1..161	
80	606	(D) OTHER INFORMATION: / Ceres Seq. ID 1941143	
		(xi) SEQUENCE DESCRIPTION: SEQ ID NO:509:	
85	607	Met Val Leu Lys Thr Glu Leu Cys Arg Phe Ser Gly Gln Lys Ile Tyr	1000 1005 1010 1015 1020 1025 1030 1035 1040 1045 1050 1055 1060 1065 1070 1075 1080 1085 1090 1095 1100 1105 1110 1115 1120 1125 1130 1135 1140 1145 1150 1155 1160 1165 1170 1175 1180 1185 1190 1195 1200 1205 1210 1215 1220 1225 1230 1235 1240 1245 1250 1255 1260 1265 1270 1275 1280 1285 1290 1295 1300 1305 1310 1315 1320 1325 1330 1335 1340 1345 1350 1355 1360 1365 1370 1375 1380 1385 1390 1395 1400 1405 1410 1415 1420 1425 1430 1435 1440 1445 1450 1455 1460 1465 1470 1475 1480 1485 1490 1495 1500 1505 1510 1515 1520 1525 1530 1535 1540 1545 1550 1555 1560 1565 1570 1575 1580 1585 1590 1595 1600 1605 1610 1615 1620 1625 1630 1635 1640 1645 1650 1655 1660 1665 1670 1675 1680 1685 1690 1695 1700 1705 1710 1715 1720 1725 1730 1735 1740 1745 1750 1755 1760 1765 1770 1775 1780 1785 1790 1795 1800 1805 1810 1815 1820 1825 1830 1835 1840 1845 1850 1855 1860 1865 1870 1875 1880 1885 1890 1895 1900 1905 1910 1915 1920 1925 1930 1935 1940 1945 1950 1955 1960 1965 1970 1975 1980 1985 1990 1995
		Pro Gly Lys Gly Ile Arg Phe Ile Arg Ala Asp Ser Gln Val Phe Leu	2000 2005 2010 2015 2020 2025 2030 2035 2040 2045 2050 2055 2060 2065 2070 2075 2080 2085 2090 2095 2100 2105 2110 2115 2120 2125 2130 2135 2140 2145 2150 2155 2160 2165 2170 2175 2180 2185 2190 2195 2200 2205 2210 2215 2220 2225 2230 2235 2240 2245 2250 2255 2260 2265 2270 2275 2280 2285 2290 2295 2300 2305 2310 2315 2320 2325 2330 2335 2340 2345 2350 2355 2360 2365 2370 2375 2380 2385 2390 2395 2400 2405 2410 2415 2420 2425 2430 2435 2440 2445 2450 2455 2460 2465 2470 2475 2480 2485 2490 2495 2500 2505 2510 2515 2520 2525 2530 2535 2540 2545 2550 2555 2560 2565 2570 2575 2580 2585 2590 2595 2600 2605 2610 2615 2620 2625 2630 2635 2640 2645 2650 2655 2660 2665 2670 2675 2680 2685 2690 2695 2700 2705 2710 2715 2720 2725 2730 2735 2740 2745 2750 2755 2760 2765 2770 2775 2780 2785 2790 2795 2800 2805 2810 2815 2820 2825 2830 2835 2840 2845 2850 2855 2860 2865 2870 2875 2880 2885 2890 2895 2900 2905 2910 2915 2920 2925 2930 2935 2940 2945 2950 2955 2960 2965 2970 2975 2980 2985 2990 2995
85	608	Phe Ala Asn Ser Lys Cys Lys Arg Tyr Phe His Asn Arg Leu Lys Pro	3000 3005 3010 3015 3020 3025 3030 3035 3040 3045 3050 3055 3060 3065 3070 3075 3080 3085 3090 3095 3100 3105 3110 3115 3120 3125 3130 3135 3140 3145 3150 3155 3160 3165 3170 3175 3180 3185 3190 3195 3200 3205 3210 3215 3220 3225 3230 3235 3240 3245 3250 3255 3260 3265 3270 3275 3280 3285 3290 3295 3300 3305 3310 3315 3320 3325 3330 3335 3340 3345 3350 3355 3360 3365 3370 3375 3380 3385 3390 3395 3400 3405 3410 3415 3420 3425 3430 3435 3440 3445 3450 3455 3460 3465 3470 3475 3480 3485 3490 3495 3500 3505 3510 3515 3520 3525 3530 3535 3540 3545 3550 3555 3560 3565 3570 3575 3580 3585 3590 3595 3600 3605 3610 3615 3620 3625 3630 3635 3640 3645 3650 3655 3660 3665 3670 3675 3680 3685 3690 3695 3700 3705 3710 3715 3720 3725 3730 3735 3740 3745 3750 3755 3760 3765 3770 3775 3780 3785 3790 3795 3800 3805 3810 3815 3820 3825 3830 3835 3840 3845 3850 3855 3860 3865 3870 3875 3880 3885 3890 3895 3900 3905 3910 3915 3920 3925 3930 3935 3940 3945 3950 3955 3960 3965 3970 3975 3980 3985 3990 3995
		(2) INFORMATION FOR SEQ ID NO:510:	
90	609	(i) SEQUENCE CHARACTERISTICS:	
		(A) LENGTH: 106 amino acids	
95	610	(B) TYPE: amino acid	
100	611	(C) STRANDEDNESS:	
		(D) TOPOLOGY: linear	
105	612	(ii) MOLECULE TYPE: peptide	
		(ix) FEATURE:	
110	613	(A) NAME/KEY: peptide	
		(B) LOCATION: 1..106	
115	614	(D) OTHER INFORMATION: / Ceres Seq. ID 1941144	
		(xi) SEQUENCE DESCRIPTION: SEQ ID NO:510:	
120	615	Met Tyr Arg Lys Gln His Lys Lys Asp Ile His Ala Glu Ala Val Lys	1000 1005 1010 1015 1020 1025 1030 1035 1040 1045 1050 1055 1060 1065 1070 1075 1080 1085 1090 1095 1100 1105 1110 1115 1120 1125 1130 1135 1140 1145 1150 1155 1160 1165 1170 1175 1180 1185 1190 1195 1200 1205 1210 1215 1220 1225 1230 1235 1240 1245 1250 1255 1260 1265 1270 1275 1280 1285 1290 1295 1300 1305 1310 1315 1320 1325 1330 1335 1340 1345 1350 1355 1360 1365 1370 1375 1380 1385 1390 1395 1400 1405 1410 1415 1420 1425 1430 1435 1440 1445 1450 1455 1460 1465 1470 1475 1480 1485 1490 1495 1500 1505 1510 1515 1520 1525 1530 1535 1540 1545 1550 1555 1560 1565 1570 1575 1580 1585 1590 1595 1600 1605 1610 1615 1620 1625 1630 1635 1640 1645 1650 1655 1660 1665 1670 1675 1680 1685 1690 1695 1700 1705 1710 1715 1720 1725 1730 1735 1740 1745 1750 1755 1760 1765 1770 1775 1780 1785 1790 1795 1800 1805 1810 1815 1820 1825 1830 1835 1840 1845 1850 1855 1860 1865 1870 1875 1880 1885 1890 1895 1900 1905 1910 1915 1920 1925 1930 1935 1940 1945 1950 1955 1960 1965 1970 1975 1980 1985 1990 1995
		Lys Arg Arg Arg Ala Thr Lys Lys Pro Tyr Ser Arg Ser Ile Val Gly	2000 2005 2010 2015 2020 2025 2030 2035 2040 2045 2050 2055 2060 2065 2070 2075 2080 2085 2090 2095 2100 2105 2110 2115 2120 2125 2130 2135 2140 2145 2150 2155 2160 2165 2170 2175 2180 2185 2190 2195 2200 2205 2210 2215 2220 2225 2230 2235 2240 2245 2250 2255 2260 2265 2270 2275 2280 2285 2290 2295 2300 2305 2310 2315 2320 2325 2330 2335 2340 2345 2350 2355 2360 2365 2370 2375 2380 2385 2390 2395 2400 2405 2410 2415 2420 2425 2430 2435 2440 2445 2450 2455 2460 2465 2470 2475 2480 2485 2490 2495 2500 2505 2510 2515 2520 2525 2530 2535 2540 2545 2550 2555 2560 2565 2570 2575 2580 2585 2590 2595 2600 2605 2610 2615 2620 2625 2630 2635 2640 2645 2650 2655 2660 2665 2670 2675 2680 2685 2690 2695 2700 2705 2710 2715 2720 2725 2730 2735 2740 2745 2750 2755 2760 2765 2770 2775 2780 2785 2790 2795 2800 2805 2810 2815 2820 2825 2830 2835 2840 2845 2850 2855 2860 2865 2870 2875 2880 2885 2890 2895 2900 2905 2910 2915 2920 2925 2930 2935 2940 2945 2950 2955 2960 2965 2970 2975 2980 2985 2990 2995
125	616	Ala Ser Leu Glu Val Ile Gln Lys Lys Arg Ala Glu Lys Pro Glu Val	3000 3005 3010 3015 3020 3025 3030 3035 3040 3045 3050 3055 3060 3065 3070 3075 3080 3085 3090 3095 3100 3105 3110 3115 3120 3125 3130 3135 3140 3145 3150 3155 3160 3165 3170 3175 3180 3185 3190 3195 3200 3205 3210 3215 3220 3225 3230 3235 3240 3245 3250 3255 3260 3265 3270 3275 3280 3285 3290 3295 3300 3305 3310 3315 3320 3325 3330 3335 3340 3345 3350 3355 3360 3365 3370 3375 3380 3385 3390 3395 3400 3405 3410 3415 3420 3425 3430 3435 3440 3445 3450 3455 3460 3465 3470 3475 3480 3485 3490 3495 3500 3505 3510 3515 3520 3525 3530 3535 3540 3545 3550 3555 3560 3565 3570 3575 3580 3585 3590 3595 3600 3605 3610 3615 3620 3625 3630 3635 3640 3645 3650 3655 3660 3665 3670 3675 3680 3685 3690 3695 3700 3705 3710 3715 3720 3725 3730 3735 3740 3745 3750 3755 3760 3765 3770 3775 3780 3785 3790 3795 3800 3805 3810 3815 3820 3825 3830 3835 3840 3845 3850 3855 3860 3865 3870 3875 3880 3885 3890 3895 3900 3905 3910 3915 3920 3925 3930 3935 3940 3945 3950 3955 3960 3965 3970 3975 3980 3985 3990 3995
		Arg Asp Ala Ala Arg Glu Ala Leu Arg Glu Ile Lys Glu Arg Ile	4000 4005 4010 4015 4020 4025 4030 4035 4040 4045 4050 4055 4060 4065 4070 4075 4080 4085 4090 4095 4100 4105 4110 4115 4120 4125 4130 4135 4140 4145 4150 4155 4160 4165 4170 4175 4180 4185 4190 4195 4200 4205 4210 4215 4220 4225 4230 4235 4240 4245 4250 4255 4260 4265 4270 4275 4280 4285 4290 4295 4300 4305 4310 4315 4320 4325 4330 4335 4340 4345 4350 4355 4360 4365 4370 4375 4380 4385 4390 4395 4400 4405 4410 4415 4420 4425 4430 4435 4440 4445 4450 4455 4460 4465 4470 4475 4480 4485 4490 4495 4500 4505 4510 4515 4520 4525 4530 4535 4540 4545 4550 4555 4560 4565 4570 4575 4580 4585 4590 4595 4600 4

5 593  
 ccgagggtt gctacgggtt tcttcgtat gttatggaga gtgggtgcaa ggggtgtgag  
 gtgacgtga gtggaaggt cagggtccaa agagcaagt ccatgaagt caagatgac  
 tacatgact catctgtca accaggaac gatacatg actcgtgtt gagacacgtt  
 cttctcagc aggtgtgtt cggatcaag gtgaagatca tcttgcagt ggaacacgtt  
 ggaaggtt ggcagcac tctcttcg gaactgtga ccatcacac ccgaaggac  
 ggaacagc cgcgcctcc ggtttggca cctcctgag tgaagttcc caactgttc  
 ctgaatgg caattatc aggtctaga tgggtttcta tcttcacata aggaacatg  
 taagcaggt gttaccta atcttatgg acagatgcta tgttgtct ctggcact  
 ttgttaag ctaagtttg cggagacag atacttat atgttccga ggaattatg  
 gttc

(2) INFORMATION FOR SEQ ID NO:512:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 274 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(i) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..274

(D) OTHER INFORMATION: / Ceres Seq. ID 1942976

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:512:

1 Ile Asn Ser Ser Ala Pro Leu Arg Arg Phe Leu Ser Arg Arg Arg Ser  
 5 10 15

25 Pro Pro Asn Pro Leu Ala Gln Asn Pro Arg Ala Leu Ala Cys Ser Pro  
 20 25 30

30 Gln Ile Ser Lys Lys Lys Phe Val Ser Asp Gly Val Phe Tyr Ala  
 35 40 45 50 55 60

65 Glu Leu Asn Glu Met Leu Thr Arg Glu Leu Ala Glu Asp Gly Tyr Ser  
 70 75 80

35 Ala Thr Arg Thr Gln Asn Val Leu Gly Glu Lys Gly Arg Arg Ile Arg  
 100 105 110 115 120 125

40 Val Glu Leu Tyr Ala Glu Lys Val Val Asn Arg Gly Leu Cys Ala Ile  
 130 135 140 145

45 Lys Gly Cys Glu Val Ile Val Ser Gly Lys Leu Arg Ala Gln Arg Ala  
 165 170 175 180 185 190 195

50 Val Asn Glu Tyr Ile Asp Ser Ala Val Arg His Val Leu Leu Arg Gln  
 210 215 220 225 230 235 240

55 Thr Pro Lys Asp Glu Asp Glu Pro Arg Pro Val Leu Ala Pro Pro  
 260 265 270

60 (2) INFORMATION FOR SEQ ID NO:513:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 229 amino acids

(B) TYPE: amino acid

594  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..229  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1942977  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:513:

1 Met Ala Thr Gln Ile Ser Lys Lys Lys Phe Val Ser Asp Gly Val  
 5 10 15

10 Phe Tyr Ala Glu Leu Asn Glu Met Leu Thr Arg Glu Leu Ala Glu Asp  
 20 25 30

15 Gly Tyr Ser Gly Val Glu Val Arg Val Thr Pro Met Arg Thr Glu Ile  
 35 40 45 50 55

60 Ile Ile Arg Ala Thr Arg Thr Gln Asn Val Leu Gly Glu Lys Gly Arg  
 65 70 75 80

20 Glu Asn Gly Val Glu Leu Tyr Ala Glu Lys Val Val Asn Arg Gly Leu  
 85 90 95

25 Cys Ala Ile Ala Gln Ala Glu Ser Leu Arg Tyr Lys Leu Leu Gly Gly  
 100 105 110 115 120 125

30 Ser Gly Ala Lys Gly Cys Glu Val Ile Val Ser Gly Lys Leu Arg Ala  
 130 135 140 145 150 155 160

35 Gln Arg Ala Lys Ser Met Lys Phe Lys Asp Gly Tyr Met Ile Ser Ser  
 165 170 175 180 185 190 195

40 Gly Gln Pro Val Asn Glu Tyr Ile Asp Ser Ala Val Arg His Val Leu  
 200 205 210 215 220

45 Leu Arg Gln Gly Val Leu Gly Ile Lys Val Lys Ile Met Leu Asp Trp  
 225 230 235 240 245 250 255

50 Thr Ile His Thr Pro Lys Asp Glu Asp Glu Pro Arg Pro Val Leu  
 260 265 270

55 Ala Pro Pro Glu Val  
 275 280 285 290 295

(2) INFORMATION FOR SEQ ID NO:514:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 206 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..206

(D) OTHER INFORMATION: / Ceres Seq. ID 1942978

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:514:

1 Met Leu Thr Arg Glu Leu Ala Glu Asp Gly Tyr Ser Gly Val Glu Val  
 5 10 15

20 Arg Val Thr Pro Met Arg Thr Glu Ile Ile Arg Ala Thr Arg Thr  
 25 30 35 40 45

55 Gln Asn Val Leu Gly Lys Gly Arg Arg Ile Arg Glu Leu Thr Ser  
 60 65 70 75 80 85 90

60 Val Val Gln Lys Arg Phe Asn Phe Pro Glu Asn Gly Val Glu Leu Tyr  
 95 100 105 110 115 120 125

120 Ala Glu Lys Val Val Asn Arg Gly Leu Cys Ala Ile Ala Gln Ala Glu  
 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195

200 Ser Leu Arg Tyr Lys Leu Leu Gly Gly Leu Ala Val Arg Arg Ala Cys  
 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295

595

Tyr Gly Val Leu Arg Tyr Val Met Glu Ser Gly Ala Lys Gly Cys Glu  
100 110  
Val Ile Val Ser Gly Lys Leu Arg Ala Gln Arg Ala Lys Ser Met Lys  
115 125  
Phe Lys Asp Gly Tyr Met Ile Ser Ser Gly Gln Pro Val Asn Glu Tyr  
130 140  
Ile Asp Ser Ala Val Arg His Val Leu Leu Arg Gln Gly Val Leu Gly  
145 155  
Ile Lys Val Lys Ile Met Leu Asp Trp Asp Pro Lys Gly Lys Val Gly  
165 175  
Pro Ile Thr Pro Leu Pro Asp Leu Val Thr Ile His Thr Pro Lys Asp  
180 190  
Glu Asp Glu Pro Arg Pro Val Leu Ala Pro Pro Glu Val  
195 205

(2) INFORMATION FOR SEQ ID NO:515:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1069 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..1069

(D) OTHER INFORMATION: / Ceres Seq. ID 1944349

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:515:

actcgtctc tgccttgcac cagctcttgc gctcggctta tctcgtcggg ttctcgttag  
aacgtctctt aactctgctt cgtgtgctgt gtttagtcat ttcatgttt tatttctgc  
attctgttgt aactctggtt cacttgatc aacataggaa gatgtacgcc cagtctgagt  
ttgactcggc tctgttgtt tccctaacta gggactgaag gctcgaact ctactcgtt  
tgcatacggg tagttctgc tctcgaagt atcagaagt cctttgagt agaaagcaca  
gcacacagat tgcacagta tggcactgt tagttagctg ttctactgac gtgtactctg  
cacagagata gacactgata ttgtgctgt ttttttggg aagattttt gaagctgaga  
agattgtgct gaagacgaaa ccttcgcgt tcaagtggca gaagatctat cctgggaag  
gcatacgggt tattcgtct gatttcagg ttctcttt tgcacactt aatcagaag  
gctcttcca caacgcctg agactctcaa agcttacctg gcagcaatg tacagaagc  
agcaagaaa ggaactccat gctgaagctg tcaagaagag gcctgcgcc accaagaagc  
catactcgaq gtcattctg gctgtctct tggaaagrat ccagaagaag agagctgaga  
agctcgaag cctgtagct gtagagaag ctgctcttcy tgaatcaag gaggcgtaca  
agagacccaa gacagagaag aagcagaaga aggtcaggt gtccaagtc cagaagcgc  
agcgaaggg cgtgtcccaag aaggttcca aggtcccaa gtgggcgcg ggcgtgaca  
aacgtgaaa agaaacgaag tctcattcg agactcaac aagctcact tctgatacc  
tgggttaatt tgcatttct gtcacaggt acttgact ctaactacc atgtaacag  
aatgattcg ttattttat atattggag aatgatcgc cgtcttcc

(2) INFORMATION FOR SEQ ID NO:516:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 161 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..161

(D) OTHER INFORMATION: / Ceres Seq. ID 1944350

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:516:

Met Val Leu Lys Thr Glu Leu Cys Arg Phe Ser Gly Gln Lys Ile Tyr  
1 15  
Pro Gly Lys Gly Ile Arg Phe Ile Arg Ala Asp Ser Gln Val Phe Leu  
20 30  
Phe Ala Asn Ser Lys Cys Lys Arg Tyr Phe His Asn Arg Leu Lys Pro  
35 45

596

Ala Lys Leu Thr Trp Thr Ala Met Tyr Arg Lys Gln His Lys Lys Asp  
50 60  
Ile His Ala Glu Ala Val Lys Lys Arg Arg Ala Thr Lys Lys Pro  
65 75  
Tyr Ser Arg Ser Ile Val Gly Ala Ser Leu Glu Val Ile Gln Lys Lys  
80 90  
Arg Ala Glu Lys Pro Glu Val Arg Asp Ala Ala Arg Glu Ala Leu  
100 110  
Arg Glu Ile Lys Glu Arg Ile Lys Lys Thr Lys Asp Glu Lys Lys Ala  
115 125  
Lys Lys Ala Glu Val Ser Lys Ser Gln Lys Ala Gln Thr Lys Gly Ala  
130 140  
Val Gln Lys Gly Ser Lys Gly Pro Lys Leu Gly Gly Gly Gly Lys  
145 155  
15 Arg 160

(2) INFORMATION FOR SEQ ID NO:517:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 106 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..106

(D) OTHER INFORMATION: / Ceres Seq. ID 1944351

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:517:

Met Tyr Arg Lys Gln His Lys Lys Asp Ile His Ala Glu Ala Val Lys  
1 15  
Lys Arg Arg Arg Ala Thr Lys Lys Pro Tyr Ser Arg Ser Ile Val Gly  
20 30  
Ala Ser Leu Glu Val Ile Gln Lys Lys Arg Ala Glu Lys Pro Glu Val  
35 40  
Arg Asp Ala Ala Arg Glu Ala Ala Leu Arg Glu Ile Lys Glu Arg Ile  
50 60  
Lys Lys Thr Lys Asp Glu Lys Lys Ala Lys Lys Ala Glu Val Ser Lys  
65 75  
Ser Gln Lys Ala Gln Thr Lys Gly Ala Val Gln Lys Gly Ser Lys Gly  
80 95  
Pro Lys Leu Gly Gly Gly Gly Lys Lys Arg  
100 105

(2) INFORMATION FOR SEQ ID NO:518:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 89 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..89

(D) OTHER INFORMATION: / Ceres Seq. ID 1944352

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:518:

Met Leu Leu Glu Lys Leu Leu Phe Val Arg Ser Arg Ser Ala Ser Arg  
1 15  
Arg Pro Arg Thr Arg Arg Arg Arg Arg Arg Arg Cys Pro Ser Pro  
20 30  
Arg Arg Arg Arg Arg Arg Ala Leu Ser Arg Arg Val Pro Ala Pro  
35 45  
Ser Trp Ala Ala Val Ala Asn Ala Glu Lys Glu Arg Ser Val Ile  
50 60

5	597	Ser Arg Pro His Gln Ser His Leu Cys Ser Thr Trp Val Asn Phe Ala	65	130	Gly Val Glu Leu Tyr Ala Glu Lys Val Val Asn Arg Gly Leu Cys Ala
		Phe Ser Phe Arg Val Thr Cys Asp Ser	85	135	Ile Ala Gln Ala Glu Ser Leu Arg Tyr Lys Leu Glu Gly Leu Ala
10	598	(2) INFORMATION FOR SEQ ID NO:519:	145	140	Val Arg Arg Ala Cys Tyr Gly Val Leu Arg Phe Val Met Glu Ser Gly
		(i) SEQUENCE CHARACTERISTICS:	150	145	Ala Glu Gly Cys Glu Val Ile Val Ser Gly Lys Leu Arg Ala Gln Arg
15	599	(A) LENGTH: 1035 base pairs	165	150	Ala Lys Ser Met Lys Phe Lys Asp Gly Tyr Met Ile Ser Ser Gly Gln
		(B) TYPE: nucleic acid	170	155	Pro Val Asn Glu Tyr Ile Asp Ser Ala Val Arg His Val Leu Leu Arg
20	600	(C) STRANDEDNESS: single	180	160	Gln Gly Val Leu Gly Ile Lys Val Ile Met Leu Asp Trp Asp Pro
		(D) TOPOLOGY: linear	185	165	Lys Gly Lys Val Gly Pro Ile Thr Pro Leu Pro Asp Leu Val Thr Ile
25	601	(ii) MOLECULE TYPE: DNA (genomic)	190	170	His Thr Pro Lys Glu Asp Glu Leu Arg Pro Pro Val Leu Ala Ala
		(ix) FEATURE:	195	175	Glu Val
30	602	(A) NAME/KEY: -	200	180	(2) INFORMATION FOR SEQ ID NO:521:
		(B) LOCATION: 1..1035	205	185	(i) SEQUENCE CHARACTERISTICS:
35	603	(D) OTHER INFORMATION: / Ceres Seq. ID 1964011	210	190	(A) LENGTH: 228 amino acids
		(xi) SEQUENCE DESCRIPTION: SEQ ID NO:519:	215	195	(B) TYPE: amino acid
40	604	aacatataa ttacgtgctc cctcttcgc gcgtctctc gcgcgcac ttccctccg	220	200	(C) STRANDEDNESS:
		accactcgc ccaaaacgtc agccctctg cctgctgcc gcgcgcgc ttgcgagctc	225	205	(D) TOPOLOGY: linear
45	605	cccatcttc tccgttaacc atggaaccc atatcagcaa gaagaaga ttgcgagctc	230	210	(ix) FEATURE:
		acggcggtt ccaacccag ataataga tgcagcgc tgcagcgc ttgcgagctc	235	215	(A) NAME/KEY: peptide
50	606	actcggctt gtagtgctt gtcacaga tgcagcgc tgcagcgc ttgcgagctc	240	220	(B) LOCATION: 1..228
		gcacagaaa cgtctcgc gagaagggc gtagatcag gtagctcag ttgcgagctc	245	225	(D) OTHER INFORMATION: / Ceres Seq. ID 1964013
55	607	gagaaggtt caattctc gagaagggc tgcagctc tgcagctc ttgcgagctc	250	230	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:521:
		ggggctctt cgcactcgc cagctcagc cctcgcga caagctcgc ttgcgagctc	255	235	Met Ala Thr Gln Ile Ser Lys Lys Lys Lys Phe Val Ser Asp Gly Val
60	608	cgtctcgc gtagctgat gtagctctt gcttgctt gtagctctt gtagctctt	260	240	1
		gtgagttat ttgagtgaa aactcaggg ctcaagagg caagctcag aagctcag	265	245	Phe Tyr Ala Glu Leu Asn Glu Met Leu Thr Arg Glu Leu Ala Glu Asp
65	609	atggtctac gatctctt gtagctcag tgaatgaga catgactca gctgtagac	270	250	Gly Tyr Ser Gly Val Glu Val Arg Val Thr Pro Met Arg Thr Glu Ile
		cgaactctt cagcagggg gtagctcag tgaatgaga catgactca gctgtagac	275	255	Ile Ile Arg Ala Thr Arg Thr Gln Asn Val Leu Gly Lys Gly Arg
70	610	aggagagaa gtagctcag atcacctc ttccgact gtagctcag catcaccca	280	260	Arg Ile Arg Glu Leu Thr Ser Val Gln Lys Arg Phe Asn Phe Pro
		tccctgaat gtagctcag cctcgcgt tgcagctc gtagctcag gtagctcag	285	265	Glu Asn Gly Val Glu Leu Tyr Ala Glu Lys Val Val Asn Arg Gly Leu
75	611	cgaagcagt gtagctcag tgcagctc gtagctcag gtagctcag gtagctcag	290	270	Cys Ala Ile Ala Gln Ala Glu Ser Leu Arg Tyr Lys Leu Leu Gly Gly
		atgactcga tgaagtgag gtagctcag gtagctcag gtagctcag gtagctcag	295	275	Leu Ala Val Arg Arg Ala Cys Tyr Gly Val Leu Arg Phe Val Met Glu
80	612	ttattttt acct	300	280	Ser Gly Ala Glu Gly Cys Glu Val Ile Val Ser Gly Lys Leu Arg Ala
		(2) INFORMATION FOR SEQ ID NO:520:	305	285	Gln Arg Ala Lys Ser Met Lys Phe Lys Asp Gly Tyr Met Ile Ser
85	613	(i) SEQUENCE CHARACTERISTICS:	310	290	145
		(A) LENGTH: 274 amino acids	315	295	Gly Gln Pro Val Asn Glu Tyr Ile Asp Ser Ala Val Arg His Val Leu
90	614	(B) TYPE: amino acid	320	300	Leu Arg Gln Gly Val Leu Gly Ile Lys Val Lys Ile Met Leu Asp Trp
		(C) STRANDEDNESS:	325	305	Asp Pro Lys Gly Lys Val Gly Pro Ile Thr Pro Leu Pro Asp Leu Val
95	615	(D) TOPOLOGY: linear	330	310	Thr Ile His Thr Pro Lys Glu Asp Glu Leu Arg Pro Pro Val Leu
		(ix) FEATURE:	335	315	Ala Ala Glu Val
100	616	(A) NAME/KEY: peptide	340	320	225
		(B) LOCATION: 1..274	345	325	
105	617	(D) OTHER INFORMATION: / Ceres Seq. ID 1964012	350	330	
		(xi) SEQUENCE DESCRIPTION: SEQ ID NO:520:	355	335	
110	618	His Ile Asn Tyr Val Leu Pro Leu Arg Arg Arg Leu Ser Arg Arg His	360	340	
		1	365	345	
115	619	Phe Pro Asp Pro Leu Ala Gln Asn Arg Ser Ala Leu Ala Cys Ser	370	350	
		20	375	355	
120	620	Pro Pro Pro Ser Leu Arg Ala Pro His Leu Pro Pro Leu Thr Met Ala	380	360	
		35	385	365	
125	621	Thr Gln Ile Ser Lys Lys Lys Phe Val Ser Asp Gly Val Phe Tyr	390	370	
		50	395	375	
130	622	Ala Glu Leu Asn Glu Met Leu Thr Arg Glu Leu Ala Glu Asp Gly Tyr	400	380	
		65	405	385	
135	623	Ser Gly Val Glu Val Arg Val Thr Pro Met Arg Thr Glu Ile Ile	410	390	
		85	415	395	
140	624	Arg Ala Thr Arg Thr Gln Asn Val Leu Gly Lys Gly Arg Arg Ile	420	400	
		100	425	405	
145	625	Arg Glu Leu Thr Ser Val Val Gln Lys Arg Phe Asn Phe Pro Glu Asn	430	410	
		115	435	415	

599

(2) INFORMATION FOR SEQ ID NO:522:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 205 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..205

(D) OTHER INFORMATION: / Ceres Seq. ID 1964014

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:522:

Met Leu Thr Arg Glu Leu Ala Glu Asp Gly Tyr Ser Gly Val Glu Val 15

1 Met Leu Thr Arg Glu Leu Ala Glu Asp Gly Tyr Ser Gly Val Glu Val 10

15 Arg Val Thr Pro Met Arg Thr Glu Ile Ile Ile Arg Ala Thr Arg Thr 25

20 Gln Asn Val Leu Gly Glu Lys Gly Arg Arg Ile Arg Glu Leu Thr Ser 30

35 Val Val Gln Lys Arg Phe Asn Phe Pro Glu Asn Gly Val Glu Leu Tyr 40

45 Ala Glu Lys Val Val Asn Arg Gly Leu Cys Ala Ile Ala Gln Ala Glu 50

55 Ser Leu Arg Tyr Lys Leu Leu Gly Gly Leu Ala Val Arg Arg Ala Cys 60

65 Tyr Gly Val Leu Arg Phe Val Met Glu Ser Gly Ala Glu Gly Cys Glu 70

80 Val Ile Val Ser Gly Lys Leu Arg Ala Gln Arg Ala Lys Ser Met Lys 85

90 Phe Lys Asp Gly Tyr Met Ile Ser Ser Gly Gln Pro Val Asn Glu Tyr 95

100 Ile Asp Ser Ala Val Arg His Val Leu Leu Arg Gln Gly Val Leu Gly 105

110 Ile Lys Val Lys Ile Met Leu Asp Trp Asp Pro Lys Gl. Lys Val Gly 115

120 Pro Ile Thr Pro Leu Pro Asp Leu Val Thr Ile His Thr Pro Lys Glu 125

130 Glu Asp Glu Leu Arg Pro Pro Val Leu Ala Ala Glu Val 135

140 (2) INFORMATION FOR SEQ ID NO:523:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 818 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..818

(D) OTHER INFORMATION: / Ceres Seq. ID 1983854

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:523:

gataaaga accctagcgg agcactaac acagtcgctt cttctctgtr cgcgcgcgcg 60

65 cgcgcgcgcg cgcgcgcgcg cgcgcgcgcg cgcgcgcgcg cgcgcgcgcg cgcgcgcgcg 120

125 agcttgtaga ggttttttga ggttgtagac atggttgtaga agacggact ttgcgccttc 180

185 agcscaga gattttctt gggaaagcga ttgatttat cgtgctgat tctcagctct 240

245 ttccttttgc caactcgaaa tgaacgctt acttcacaa cgcctctgaag cgtgcgaagc 300

305 ttcactgac agcaatgac aggaagcagc acaagaaga tatctatgt gaagcgttaa 360

365 aagaagcgc cgcgcgcac agaaagccat actccagtc atttgttgt gctctctgg 420

425 agcaatcca gaagaaga gctgagaagc cagagctccg cgtgctgct agagaagctg 480

485 cttctctga gatcaaggag cgtatcaaga agaccaaga tgaagaaga gcgaagaagg 540

545 cggagggag agactccag agactccag caaaggctg gctgagaag gtttccaag 600

605 gccccaagt gggcgccgt ggtggagag gctgagaag gctgagctg ttctcgaa 660

720 ttgcagcgt tctctgaca agccacttt cgtgaacct cgttgactt tgaagactt 780

785

790

795

800

805

810

815

820

825

830

835

840

845

850

855

860

865

870

875

880

885

890

895

600

(2) INFORMATION FOR SEQ ID NO:524:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 106 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..106

(D) OTHER INFORMATION: / Ceres Seq. ID 1983855

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:524:

Met Tyr Arg Lys Gln His Lys Lys Asp Ile His Ala Glu Ala Val Lys 15

1 Met Tyr Arg Lys Gln His Lys Lys Asp Ile His Ala Glu Ala Val Lys 10

15 Lys Arg Arg Arg Ala Thr Lys Lys Pro Tyr Ser Arg Ser Ile Val Gly 20

25 Ala Ser Leu Glu Val Ile Gln Lys Lys Arg Ala Glu Lys Pro Glu Val 30

35 Arg Asp Ala Ala Arg Glu Ala Ala Leu Arg Glu Ile Lys Glu Arg Ile 40

45 Lys Lys Thr Lys Asp Glu Lys Lys Ala Lys Lys Lys Ala Glu Val Ser Lys 50

55 Ser Gln Lys Thr Gln Thr Lys Gly Ala Val Gln Lys Gly Ser Lys Gly 60

65 Pro Lys Leu Gly Gly Gly Gly Lys Arg 70

75 (2) INFORMATION FOR SEQ ID NO:525:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 802 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..802

(D) OTHER INFORMATION: / Ceres Seq. ID 1990261

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:525:

gataaaga accctagcgg mscaataaya cagtcgctc cttctctgtr cgcgcgcgcg 60

65 ggcgcgcgcg cgcgcgcgcg ggcgcgcgcg ggcgcgcgcg ggcgcgcgcg ggcgcgcgcg 120

125 gaagtttct tgagctgag aacatgggc tgaagaagg actttgcgc ttcagcgccg 180

185 agaagattta tctgggaaa ggcattgat ttatcgtgc tgattctcg gttctcttt 240

245 ttgcacactt gaatgcga cgtacttcc acaaccgct gaacttcca ggcgcgcgcg 300

305 ggcgcgcgcg caccagaag ccaactcca ggcattgct ggcgcgcgcg tgaagaaga 360

365 tccagaaga gagactgag aagccaggg tccgcgatg tgcctagaa gctgcttcc 420

425 ggcgcgcgcg gagactgag aagccaggg tccgcgatg tgcctagaa gctgcttcc 480

485 tgaagcagtc ccagaagcg cagacaagg gtcgcgcgcg ggcgcgcgcg aagcgagag 540

545 agttggcgcg cgtgtgtgg aagcctgaa agacttagt gtcgttctc gacattgag 600

605 tgcgttctc gcaagtcga cttctgaga accgtgtgt aatttgaag acttattcaa 660

665 aaaaatgacc gttcattatc tt 720

725 (2) INFORMATION FOR SEQ ID NO:526:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 161 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..161

(D) OTHER INFORMATION: / Ceres Seq. ID 1990261

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:526:

gataaaga accctagcgg mscaataaya cagtcgctc cttctctgtr cgcgcgcgcg 60

65 ggcgcgcgcg cgcgcgcgcg ggcgcgcgcg ggcgcgcgcg ggcgcgcgcg ggcgcgcgcg 120

125 gaagtttct tgagctgag aacatgggc tgaagaagg actttgcgc ttcagcgccg 180

185 agaagattta tctgggaaa ggcattgat ttatcgtgc tgattctcg gttctcttt 240

245 ttgcacactt gaatgcga cgtacttcc acaaccgct gaacttcca ggcgcgcgcg 300

305 ggcgcgcgcg caccagaag ccaactcca ggcattgct ggcgcgcgcg tgaagaaga 360

365 tccagaaga gagactgag aagccaggg tccgcgatg tgcctagaa gctgcttcc 420

425 ggcgcgcgcg gagactgag aagccaggg tccgcgatg tgcctagaa gctgcttcc 480

485 tgaagcagtc ccagaagcg cagacaagg gtcgcgcgcg ggcgcgcgcg aagcgagag 540

545 agttggcgcg cgtgtgtgg aagcctgaa agacttagt gtcgttctc gacattgag 600

605 tgcgttctc gcaagtcga cttctgaga accgtgtgt aatttgaag acttattcaa 660

665 aaaaatgacc gttcattatc tt 720

725 (2) INFORMATION FOR SEQ ID NO:527:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 161 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

601

(A) NAME/KEY: peptide  
(B) LOCATION: 1..161  
(D) OTHER INFORMATION: / Ceres Seq. ID 2061972  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:526:

5 Met Val Leu Lys Thr Glu Leu Cys Arg Phe Ser Gly Gln Lys Ile Tyr  
1 5 10 15  
Pro Gly Lys Gly Ile Arg Phe Ile Arg Ala Asp Ser Gln Val Phe Leu  
20 25 30  
Phe Ala Asn Ser Lys Cys Lys Arg Tyr Phe His Asn Arg Leu Lys Pro  
35 40 45  
Ala Lys Leu Thr Trp Thr Ala Met Tyr Arg Arg Lys Gln His Lys Lys Asp  
50 55 60  
Ile His Ala Glu Ala Val Lys Lys Arg Arg Arg Ala Thr Lys Lys Pro  
65 70 75 80  
Tyr Ser Arg Ser Ile Val Gly Ala Ser Leu Glu Val Ile Gln Lys Lys  
90 95  
Arg Ala Glu Lys Pro Glu Val Arg Asp Ala Ala Arg Glu Ala Ala Leu  
100 105 110  
Arg Glu Ile Lys Glu Arg Ile Lys Lys Thr Lys Asp Glu Lys Lys Ala  
115 120 125  
Lys Lys Ala Glu Val Ser Lys Ser Gln Lys Thr Gln Thr Lys Gly Ala  
130 135 140  
Val Gln Lys Gly Ser Lys Gly Pro Lys Leu Gly Gly Gly Gly Lys  
145 150 155 160  
Arg

(2) INFORMATION FOR SEQ ID NO:527:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 106 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..106  
(D) OTHER INFORMATION: / Ceres Seq. ID 2061973  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:527:

40 Met Tyr Arg Lys Gln His Lys Lys Asp Ile His Ala Glu Ala Val Lys  
1 5 10 15  
Lys Arg Arg Arg Ala Thr Lys Lys Pro Tyr Ser Arg Ser Ile Val Gly  
20 25 30  
Ala Ser Leu Glu Val Ile Gln Lys Lys Arg Ala Glu Lys Pro Glu Val  
35 40 45  
Arg Asp Ala Ala Arg Glu Ala Ala Leu Arg Glu Ile Lys Glu Arg Ile  
50 55 60  
Lys Lys Thr Lys Asp Glu Lys Lys Ala Lys Lys Ala Glu Val Ser Lys  
65 70 75 80  
Ser Gln Lys Thr Gln Thr Lys Gly Ala Val Gln Lys Gly Ser Lys Gly  
85 90 95  
Pro Lys Leu Gly Gly Gly Gly Lys Arg  
100 105

## SEQ TABLE 2

(2) INFORMATION FOR SEQ ID NO:1:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 494 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

602

(A) NAME/KEY: -  
(B) LOCATION: 1..494  
(D) OTHER INFORMATION: / Ceres Seq. ID 1007802  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

5 aacacacgga agaaaggaga agacaaagat gtcgttggtg tgggtggaag ccatgttgcc  
1 tctcgaatc atcggtggga tctcttgat catgggcaat tctcagtact acatcacaaa  
6 agttatcat gccctccta agcacatcg ccaagargaa tgggagtgtg ctatgggaag  
12 acggacaaag aaagctgcg agaaagctgc agctctctcc tcatgattcg cttattctct  
18 ttgtctcc t-aggggctt aagtgacca ctgtgtgta caataaagat gcattccaga  
24 agaaagaa gctgggggat ctagtactt cattccatt tgcattccc tggacattt  
30 aaagcttca gaatacagac ctaataaca ttgggtttat caattttct ccatctgga  
36 caatttgat gcttttaagt tgcargatac actgaactac tctctgtgtg tgrgagtga  
42 caatggaata tctc  
48

(2) INFORMATION FOR SEQ ID NO:2:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 94 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: Peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..94  
(D) OTHER INFORMATION: / Ceres Seq. ID 1007803  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

1 Gln Ser Glu Glu Arg Arg Arg Arg Arg Cys Arg Trp Tyr Gly Trp Lys  
5 10 15  
Arg Cys Cys Leu Ser Glu Ser Ser Val Gly Cys Ser Val Ser Trp Ala  
20 25 30  
Ile Leu Ser Thr Thr Ser Thr Lys Leu Ile Met Ala Val Leu Ser Thr  
35 40 45  
Ser Ala Thr Met Asn Gly Met Leu Leu Trp Lys Asp Ala Thr Arg Lys  
50 55 60  
Ser Ser Arg Lys Leu Gln Leu Leu Pro His Asp Ser Leu Tyr Leu Phe  
65 70 75 80  
Cys Val Pro Gln Gly Leu Lys Val Thr Thr Cys Gly Asp Lys  
85

(2) INFORMATION FOR SEQ ID NO:3:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 74 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..74  
(D) OTHER INFORMATION: / Ceres Seq. ID 1007804  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

1 Thr Ile Gly Arg Lys Glu Lys Thr Lys Met Ser Leu Val Trp Leu Glu  
5 10 15  
Ala Met Leu Pro Leu Gly Ile Ile Gly Gly Met Leu Cys Ile Met Gly  
20 25 30  
Asn Ser Gln Tyr Tyr Ile His Lys Ala Tyr His Gly Arg Pro Lys His  
35 40 45  
Ile Gly His Asp Glu Trp Asp Val Ala Met Glu Arg Arg Asp Lys Lys  
50 55 60  
Val Val Glu Lys Ala Ala Pro Ser Ser  
65 70

(2) INFORMATION FOR SEQ ID NO:4:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 65 amino acids

603

(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide  
(B) LOCATION: 1..65

(D) OTHER INFORMATION: / Ceres Seq. ID 1007805  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Ser Leu Val Trp Leu Glu Ala Met Leu Pro Leu Gly Ile Ile Gly  
1 5 10 15

Gly Met Leu Cys Ile Met Gly Asn Ser Gln Tyr Tyr Ile His Lys Ala  
20 25 30

Tyr His Gly Arg Pro Lys His Ile Gly His Asp Glu Trp Asp Val Ala  
35 40 45

Met Glu Arg Arg Asp Lys Lys Val Val Glu Lys Ala Ala Pro Ser  
50 55 60

Ser

(2) INFORMATION FOR SEQ ID NO:5:  
(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 700 base pairs  
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..700

(D) OTHER INFORMATION: / Ceres Seq. ID 1008556  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

aaagaagaaac tgaagaaacc caaaataccc aagtgaaaga aacgtttatc cacttcgaaga  
gagtgacacg tgcagaaaga aatcaaat .aaactttt caataaaat caaactcatt  
cttttatta cttcccacaa tgcgaatacag taacttatta aaacctataa tcatgacatt  
cgacccgaca tgcagacaga cgttcgtgc atcaacaaac ttcctgaac attctttc  
cgagtgata caatacaca acatagatgat tcttccaaa acaagagat ctctacctc  
agcactccc gtgcactcg tgcgaactgc ttctgcaga acaattact taatggcgg  
gtctcggaa tctgatggag cagatattgt gaggaagaaga ggaagtgaac tgaagacca  
cagccgacgg aaggaatttc acgatcgtt tigtctaat tgcgaaggac tgcgaagct  
taaccagatt gataaatcat aagctttctt catatalatg taaaagctc cactttcat  
tctctttt attacaata tactcttcta agatagata gcttgaatg atcaatggtt  
gcttgagat gaaatgatat atatatcaa ttgcttttt  
60 120 180 240 300 360 420 480 540 600 660

(2) INFORMATION FOR SEQ ID NO:6:  
(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 129 amino acids  
(B) TYPE: amino acid

(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide  
(B) LOCATION: 1..129

(D) OTHER INFORMATION: / Ceres Seq. ID 1008557  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Thr Ile Ala Pro Ala Leu Gln Thr Thr Phe Val Ser Ser Thr Asn  
1 5 10 15

Phe Leu Lys His Ser Ser Trp Gly Ser Ser Ser Pro Asn Asn Val  
20 25 30

Ile Leu Pro Lys Asn Lys Arg Ser Ser Thr Ser Val Val Ala Ala  
35 40 45

Val Gly Asp Val Ser Ser Asp Gly Thr Ile Tyr Leu Ile Gly Gly Ala

604

(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide  
(B) LOCATION: 1..65

(D) OTHER INFORMATION: / Ceres Seq. ID 1008628  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Ala Ser Leu Val Ala Ala Pro Ile Ser Phe Ser Gly Asp Ser His  
1 5 10 15

Val Lys Ala His Arg Asn Phe Asn Ala Ile Arg Lys Ser Ser Thr Leu  
20 25 30

Thr Val Gln Thr Lys Ser Asn Arg Ser His Lys Leu Ser Val Ser Ala  
35 40 45

Gly Tyr Arg Gly Gly Ser Lys Gly Gly Ser Ser Asp Phe Val Thr  
50 55 60

Gly Phe Leu Leu Gly Ser Ala Val Phe Gly Thr Leu Ala Tyr Ile Phe  
65 70 75 80

Ala Pro Gln Ile Arg Arg Ser Val Leu Ser Glu Asn Glu Tyr Gly Phe  
85 90 95

Lys Lys Pro Glu Gln Pro Met Tyr Tyr Asp Glu Gly Leu Glu Arg  
100 105 110

Arg Glu Ile Leu Asn Glu Lys Ile Gly Gln Leu Asn Ser Ala Ile Asp  
115 120 125

Ser

(2) INFORMATION FOR SEQ ID NO:7:  
(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 665 base pairs  
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..665

(D) OTHER INFORMATION: / Ceres Seq. ID 1008628  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

atgttacct tctcagatct cttctcgtt ttgatgaca atggcgctct tggtagcagc  
tcttacct tctcagtg actctcatgt caagcacac caaacttca atgcgattcg  
caagatctt acattgactg ttcaacaaa atcaaacgc atcaaacac tctcggttc  
tgcagttac cgtggggaa gaaagggtg tgaagtatg gatttgta cggttttc  
tctagaagt gctgtctg gaacttgc ttatactt gctccaga tccgaagtc  
agtcagc gagatgaat atgtttcaa gaaccggag cagccgatg actatgaaga  
agcctagag gagagaag agatattgaa tgagaaatc gccaactca atccgcct  
tcacaagtt tgcgcgnc tgaaggagg tggagcgtg agcagaaga acactttc  
gcgcgtgc ccattgaa ccgacgaga agcaagct actgatgat tgaatgaat  
ctctgctt atttaca ttcaactg cttcattg gttctggt tttttttt  
ttttgttg aaccattagg ggtttttctg acttgaat attgaaga aaagaacac  
60 120 180 240 300 360 420 480 540 600 660

(2) INFORMATION FOR SEQ ID NO:8:  
(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 162 amino acids  
(B) TYPE: amino acid

(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

(A) NAME/KEY: peptide  
(B) LOCATION: 1..162

(D) OTHER INFORMATION: / Ceres Seq. ID 1008629  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Ala Ser Leu Val Ala Ala Pro Ile Ser Phe Ser Gly Asp Ser His  
1 5 10 15

Val Lys Ala His Arg Asn Phe Asn Ala Ile Arg Lys Ser Ser Thr Leu  
20 25 30

Thr Val Gln Thr Lys Ser Asn Arg Ser His Lys Leu Ser Val Ser Ala  
35 40 45

Gly Tyr Arg Gly Gly Ser Lys Gly Gly Ser Ser Asp Phe Val Thr  
50 55 60

Gly Phe Leu Leu Gly Ser Ala Val Phe Gly Thr Leu Ala Tyr Ile Phe  
65 70 75 80

Ala Pro Gln Ile Arg Arg Ser Val Leu Ser Glu Asn Glu Tyr Gly Phe  
85 90 95

Lys Lys Pro Glu Gln Pro Met Tyr Tyr Asp Glu Gly Leu Glu Arg  
100 105 110

Arg Glu Ile Leu Asn Glu Lys Ile Gly Gln Leu Asn Ser Ala Ile Asp

115 Lys Val Ser Ser Arg Leu Lys Gly Arg Ser Gly Ser Lys Asn 125  
 130 Thr Ser Ser Pro Val Pro Val Glu Thr Asp Ala Glu Ala 140  
 145 Thr Ala 150 155 160  
 606  
 Tyr Ile Ile His Leu Phe Phe Cys Ile Met Tyr Gly Tyr Gly Trp Pro Ser  
 100 105 110  
 Ser Met Ala Trp Trp Val Val Asn Gly Thr Gly Leu Ala Val Met Ala  
 115 120 125  
 5 Leu Leu Ala Glu Tyr Leu Cys Ile Lys Arg Glu Gln Arg Glu Ile Pro  
 130 135 140  
 Met Asp Arg Phe His Ser Arg Val 150  
 145 150  
 (2) INFORMATION FOR SEQ ID NO:11:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 120 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..120  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1009372  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:  
 Met Val Phe Leu Gly Leu Arg Val Pro Arg Leu Ser Leu Val Tyr Phe  
 1 10 15  
 Phe Asp Tyr Ala Thr Leu Thr Thr Ser Thr Phe Thr Gly Trp Ser Val  
 20 25 30  
 25 Ile Ala Ser Phe Leu Phe Ser Ser Leu Ala Gly Ala Val Tyr Met Ile  
 35 40 45  
 Phe Leu Val Glu Arg Ala Arg Lys Cys Leu Asp Phe Ser Ala Thr Leu  
 50 55 60  
 Tyr Ile Ile His Leu Phe Phe Cys Ile Met Tyr Gly Tyr Trp Pro Ser  
 65 70 75 80  
 Ser Met Ala Trp Trp Val Val Asn Gly Thr Gly Leu Ala Val Met Ala  
 85 90 95  
 Leu Leu Ala Glu Tyr Leu Cys Ile Lys Arg Glu Gln Arg Glu Ile Pro  
 100 105 110  
 Met Asp Arg Phe His Ser Arg Val 120  
 115 120  
 (2) INFORMATION FOR SEQ ID NO:12:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 76 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..76  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1009379  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:  
 Met Ala Gln Leu Tyr Gly Thr Leu Gly Leu Ser Leu Ala Arg Leu Phe  
 1 5 10 15  
 Ala Ser Asn Ala Leu Thr Ile Ser Leu Leu Asp Ser Pro Trp Ser  
 20 25 30  
 Phe Leu Ala Phe Val Phe Leu Ala Leu Val Leu Ser Thr Ser Ile  
 35 40 45  
 55 Thr Leu Leu Ser Leu Pro Pro Ser Pro Val Gly Leu Leu Pro  
 50 55 60  
 His Ser Ser Ser Leu His Ser Leu Gly Leu Cys Thr  
 65 70 75  
 (2) INFORMATION FOR SEQ ID NO:13:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 398 base pairs  
 (B) TYPE: nucleic acid

60 cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt  
 120 gttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 180 gttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 240 cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt  
 300 tttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 360 cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt  
 420 tttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 480 gttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 540 cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt  
 600 gttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 660 cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt  
 720 gttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 780 cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt  
 840 gttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 900 cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt  
 960 gttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 1020 cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt  
 1080 tttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 606  
 Lys Val Ser Ser Arg Leu Lys Gly Arg Ser Gly Ser Lys Asn 125  
 130 Thr Ser Ser Pro Val Pro Val Glu Thr Asp Ala Glu Ala 140  
 145 Thr Ala 150 155 160  
 606  
 Tyr Ile Ile His Leu Phe Phe Cys Ile Met Tyr Gly Tyr Gly Trp Pro Ser  
 100 105 110  
 Ser Met Ala Trp Trp Val Val Asn Gly Thr Gly Leu Ala Val Met Ala  
 115 120 125  
 5 Leu Leu Ala Glu Tyr Leu Cys Ile Lys Arg Glu Gln Arg Glu Ile Pro  
 130 135 140  
 Met Asp Arg Phe His Ser Arg Val 150  
 145 150  
 (2) INFORMATION FOR SEQ ID NO:9:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1101 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..1101  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1009376  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:  
 cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt  
 120 gttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 180 gttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 240 cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt  
 300 tttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 360 cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt  
 420 tttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 480 gttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 540 cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt  
 600 gttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 660 cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt  
 720 gttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 780 cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt  
 840 gttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 900 cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt  
 960 gttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 1020 cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt cttcttctt  
 1080 tttttttt tttttttt tttttttt tttttttt tttttttt tttttttt  
 606  
 Lys Val Ser Ser Arg Leu Lys Gly Arg Ser Gly Ser Lys Asn 125  
 130 Thr Ser Ser Pro Val Pro Val Glu Thr Asp Ala Glu Ala 140  
 145 Thr Ala 150 155 160  
 606  
 Tyr Ile Ile His Leu Phe Phe Cys Ile Met Tyr Gly Tyr Gly Trp Pro Ser  
 100 105 110  
 Ser Met Ala Trp Trp Val Val Asn Gly Thr Gly Leu Ala Val Met Ala  
 115 120 125  
 5 Leu Leu Ala Glu Tyr Leu Cys Ile Lys Arg Glu Gln Arg Glu Ile Pro  
 130 135 140  
 Met Asp Arg Phe His Ser Arg Val 150  
 145 150  
 (2) INFORMATION FOR SEQ ID NO:10:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 152 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..152  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1009377  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:  
 Met Phe Tyr Gly Thr Ala Val Trp Asp Pro Trp Leu Ile Val Gly Gln  
 1 5 10 15  
 Ile Ile Cys Leu Gln Cys Ser Tyr Tyr Leu Thr Thr Leu Gly Leu Phe Thr  
 20 25 30  
 Met Val Phe Leu Gly Leu Arg Val Pro Arg Leu Ser Leu Val Tyr Phe  
 35 40 45  
 Phe Asp Tyr Ala Thr Leu Thr Thr Ser Thr Thr Phe Thr Gly Trp Ser Val  
 50 55 60  
 Ile Ala Ser Phe Leu Phe Ser Ser Leu Ala Gly Ala Val Tyr Met Ile  
 65 70 75 80  
 Phe Leu Val Glu Arg Ala Arg Lys Cys Leu Asp Phe Ser Ala Thr Leu  
 85 90 95



607

- (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

5

- (A) NAME/KEY: -  
(B) LOCATION: 1..396  
(D) OTHER INFORMATION: / Ceres Seq. ID 1011128

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

10 aatcacaatg ctatcctagt ttctgactgt tcaacgatca atggcaccat cgcctgcaat 60  
gctcacaatc tagttagctc tagttagcga caaagctaaag atcacaatc ttctatgcgc 120  
ttctgattgt aagtcgacac agtctctcgg ttctcttgg ccttgaagg catatgacaa 180  
cgagatcat tccgggttg ttcaagcgc gctcttggc gaccgggcta ctatcgagaa 240  
gccttccaa gaagctcttg acaaaagctg ttgtaatt ttacaattg ttgttgggtt 300  
tattgattg ttctagctta aaatcgctt ttataaattgg aaatgaaagt actgtaaat 360

- (2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 78 amino acids  
(B) TYPE: amino acid

- (C) STRANDEDNESS:

- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: peptide

- (ix) FEATURE:

- (A) NAME/KEY: peptide

- (B) LOCATION: 1..78

- (D) OTHER INFORMATION: / Ceres Seq. ID 1011129

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Ala Pro Ser Ala Ala Met Leu Ile Leu Ser His Pro Leu Val Ser 1  
1 His Lys Ala Lys Asn Gln Ser Leu Ser Ser Pro Ser Ser Val Lys Ser 5  
20 Thr Arg Val Phe Gly Phe Leu Trp Trp Lys Ala Leu Asp Asn Gln 15  
35 Asp His Ser Ala Val Val Leu Gly Arg Leu Phe Gly Asp Pro Ala Thr 40  
50 Ile Glu Lys Arg Phe Gln Glu Ala Leu Glu Gln Ser Cys Trp 45  
65

- (2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 72 amino acids

- (B) TYPE: amino acid

- (C) STRANDEDNESS:

- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: peptide

- (ix) FEATURE:

- (A) NAME/KEY: peptide

- (B) LOCATION: 1..72

- (D) OTHER INFORMATION: / Ceres Seq. ID 1011130

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met Leu Ile Leu Ser His Pro Leu Val Ser His Lys Ala Lys Asn Gln 1  
1 Ser Leu Ser Ser Pro Ser Val Lys Ser Thr Arg Val Phe Gly Phe 10  
20 Leu Trp Pro Trp Lys Ala Leu Asp Asn Glu Asp His Ser Ala Val Val 15  
35 Leu Gly Arg Leu Phe Gly Asp Pro Ala Thr Ile Glu Lys Arg Phe Gln 40  
50 Glu Ala Leu Glu Gln Ser Cys Trp 45  
65

- (2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 549 base pairs

608

- (B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:

5

- (A) NAME/KEY: -  
(B) LOCATION: 1..549  
(D) OTHER INFORMATION: / Ceres Seq. ID 1011718

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

10 agggatgtag taggccaata ggcaatcag agaatccaaa atggtatctg gtcaagaaga 60  
ttctgtttg ttgacagatg catgtgtta atctagggtg tatgttttg tccatttgt 120  
ttcataagc aataaagtc cagctattta ctactgtga agaaaaagt gagggaaac 180  
agaacaaat ccgataact ttctggaaa gctgaggtt ttgtttctt agagacaga 240  
tgaagaggtg actgtttgtt ctgtcgagat agaatgggc actcaattt gctttcttg 300  
gaccagtga atcaattgca gctctctata tatttagta ggcgatgtat caatggtaa 360  
gcgagatga catctgtggg tactgttct taactctca ccttgatc tatgcttag 420  
aagcgttca agtctgaa gcgatttga ttgagttt ttaacaaa gaatcaattc 480  
actccaaaa tatgtaaat ctgcatgcc ttgagctgt tttaggatta tccactgtt 540  
tgttttgt

- (2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 70 amino acids

- (B) TYPE: amino acid

- (C) STRANDEDNESS:

- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: peptide

- (ix) FEATURE:

- (A) NAME/KEY: peptide

- (B) LOCATION: 1..70

- (D) OTHER INFORMATION: / Ceres Seq. ID 1011719

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Gly Cys Ser Arg Pro Ile Gly Lys Ser Glu Asn His Lys Trp Tyr Leu 1  
1 Val Lys Lys Ile Pro Val Cys Leu Gln Met His Val Cys Asn Leu Gly 5  
20 Tyr Met Phe Leu Ser Ile Trp Phe His Lys Ala Ile Lys Ile Gln Leu 10  
35 Phe Thr Thr Cys Glu Glu Lys Ser Glu Gly Lys Gln Ser Gln Val Arg 15  
50 Leu Leu Phe Trp Glu Ser 20  
65

- (2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 amino acids

- (B) TYPE: amino acid

- (C) STRANDEDNESS:

- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: peptide

- (ix) FEATURE:

- (A) NAME/KEY: peptide

- (B) LOCATION: 1..45

- (D) OTHER INFORMATION: / Ceres Seq. ID 1011720

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met His Val Cys Asn Leu Gly Tyr Met Phe Leu Ser Ile Trp Phe His 1  
1 Lys Ala Ile Lys Ile Gln Leu Phe Thr Thr Cys Glu Glu Lys Ser Glu 5  
20 Gly Lys Gln Ser Gln Val Arg Leu Leu Phe Trp Glu Ser 10  
35

- (2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 amino acids

609

(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..37  
(D) OTHER INFORMATION: / Ceres Seq. ID 1011721  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:  
1 Met Phe Leu Ser Ile Trp Phe His Lys Ala Ile Lys Ile Gln Leu Phe  
10 Thr Thr Cys Glu Glu Lys Ser Glu Gly Lys Gln Ser Gln Val Arg Leu  
20 Leu Phe Trp Glu Ser  
15 (2) INFORMATION FOR SEQ ID NO:20:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 417 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..417  
(D) OTHER INFORMATION: / Ceres Seq. ID 1011735  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:  
aattctctt ttttcaatc taccataat tccaaacac catcgattt ttgtctctt  
gaatgctt ttttatgac aattcttc ttccaactaa accgaacgg aaaccaggt  
tcggagatcg ggtttactg atggcaaac agcaacgaac cgcactttac atactcgaa  
gaigcttc catgctgct ttgtgacgc accactctat ttctgattag ttgcagtaa  
tgaaccggac tgacaagaa tgacaacta gaagaatg agtcaattca gacttttc  
ttgtttctt tgaagactt cttaaggaa tctttgaa ttgaagaaga ctctaagat  
tggcctttt gggcctttt acaratgtg taataaact atgaacttt tttaacc  
35 (2) INFORMATION FOR SEQ ID NO:21:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 51 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..51  
(D) OTHER INFORMATION: / Ceres Seq. ID 1011736  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:  
Met Ser Asn Ser Ser Leu Pro Thr Lys Pro Asn Arg Lys Thr Arg Phe  
1 Gly Asp Arg Cys Leu Leu Met Ala Lys Gln Arg Thr Arg Leu Tyr  
10 Ile Leu Arg Arg Cys Val Ser Met Leu Leu Cys Trp His Asp His Ser  
15 Ile Ser Asp  
20 (2) INFORMATION FOR SEQ ID NO:22:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 46 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

610

(A) NAME/KEY: peptide  
(B) LOCATION: 1..46  
(D) OTHER INFORMATION: / Ceres Seq. ID 1011737  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:  
5 Asn Ser Phe Phe Thr Thr Gln Asn Ser Lys Thr Pro Ser Ile  
10 Phe Cys Ser Leu Glu Met Leu Phe Leu Cys Pro Ile Leu Leu Phe Gln  
20 Leu Asn Arg Thr Gly Lys Pro Gly Ser Glu Ile Gly Val Tyr  
35 (2) INFORMATION FOR SEQ ID NO:23:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 32 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..32  
(D) OTHER INFORMATION: / Ceres Seq. ID 1011738  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:  
1 As Pro Ser Phe Ser His Leu Pro Lys Ile Pro Lys His His Arg Phe  
5 Phe Ala Leu Leu Lys Cys Ser Phe Tyr Val Gln Phe Phe Ser Ser Asn  
10 (2) INFORMATION FOR SEQ ID NO:24:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 712 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..712  
(D) OTHER INFORMATION: / Ceres Seq. ID 1011755  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:  
acctaaatag atgaagaat ggcagctta agcacagct ttctgggaat aagacgagg 60  
tcaagtaaca agacaacagg agagataag cccaataacc agtgtttgc aagaagcacc 120  
caaagcatag acaatcacct ggtgttgtt ctcttgtct taacgagagg ctctctctt 180  
ttataaagg agctcttca cgtagacca gatctgtca aacttgtct ctctctctt 240  
caactacttc ctctcttca tctacaggt ctctctgtg ctctcttgt ctctctctt 300  
ttgttgatcg ccgacgatac ttgttaagt ccggagcgag cgtagagga gagaagtga 360  
ttctggagt ccgaaaaagc cgtacctag ctataaagc ggtatgag aagaagaga 420  
agaagaagc aaagacgaat agtgtttct ttgtgtttt ggtaatggc aaaaagaaga 480  
gacaataga tagtactct tgcagattg ttgattaga ttgttttga ttgttattt 540  
tcctgtctt ttcatttag atgattgct acatagat gactagatca gtttggtgg 600  
taagtgcag tgagagagat tcttatacg tattctgtg atgattccc taagtctga 660  
gttagcttg gaacataaa atgtgataat ccaatttaa tatggttga ac

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 108 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..108  
(D) OTHER INFORMATION: / Ceres Seq. ID 1011756  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

- 611  
 Thr Gln Ile Asp Glu Glu Met Ala Ser Leu Ser Thr Ser Leu Leu Gly  
 1 5 10 15  
 Ile Arg Arg Gly Ser Ser Asn Lys Asn Lys Ala Glu Asp Lys Thr Lys  
 20 25 30  
 5 Asn Gln Cys Phe Ala Arg Ser Ile Gln Ser Ile Asp Asn His Leu Val  
 35 40 45  
 Phe Val Leu Phe Val Leu Thr Arg Gly Ser Leu Phe Leu Ser Lys Gln  
 50 55 60  
 Leu Leu His Val Asp Gln Asp Leu Val Lys Ser Cys Leu Leu Pro  
 65 70 75 80  
 Gln Leu Leu Pro Leu Phe His Leu Thr Val Leu Leu Ser Leu Leu  
 85 90 95  
 Val Leu Leu Leu Leu Ile Ala Asp Thr Cys  
 100 105
- 15 (2) INFORMATION FOR SEQ ID NO:26:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 102 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ix) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..102  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011757
- 20 Met Ala Ser Leu Ser Thr Ser Leu Leu Gly Ile Arg Arg Gly Ser Ser  
 1 5 10 15  
 Ser Ile Gln Ser Ile Asp Asn His Leu Val Phe Val Leu Phe Val Leu  
 20 25 30  
 Thr Arg Gly Ser Leu Phe Leu Ser Lys Gln Leu Leu His Val Asp Gln  
 35 40 45  
 Asp Leu Val Lys Ser Cys Leu Leu Leu Pro Gln Leu Leu Pro Leu Phe  
 50 55 60  
 His Leu Thr Val Leu Leu Ser Leu Leu Val Leu Leu Leu Leu  
 65 70 75 80  
 Ile Ala Asp Thr Cys  
 85 90 95
- 40 (2) INFORMATION FOR SEQ ID NO:27:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 422 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ix) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..422  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011832
- 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:  
 aataacaata tcaaaagata aaacagagt tgcgttttat aagatcaaaag aaggtaaaaa  
 60  
 aaaaagtg9 caaatggag tgcacgcga ttgataatga tggtagatg tggtagcgc  
 120  
 acaatagaag cacaagaaga aagtgattgg actattgtt ttcccaatg ttctcagct  
 180  
 tgcgcagt9 acgcagtga ttgtatgaa aacttaaaa tcgaatcgg tggcccaag  
 240  
 ctcccttat ctactcac9 aagtcacat gcataagatg gcatacaca tggaaattcg  
 300  
 cgaagga9 gacagataa aagagattg tgcggattt ttgcatact ctactgttt  
 360  
 taactcac9 tttatata taadatttc attaaaaga tcaataaac gagaaatgt  
 420
- 50 (2) INFORMATION FOR SEQ ID NO:28:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 713 base pairs  
 (B) TYPE: nucleic acid
- 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:  
 Met Ile Val Ala Val Thr Ile Glu Ala Gln Glu Ser Gly Trp Thr  
 1 5 10 15  
 Ile Cys Phe Arg Gln Cys Ser Gln Pro Cys Arg Ser Asp Gly Ser  
 20 25 30  
 Cys Tyr Glu Asn Cys Lys Ile Glu Cys Gly Pro Lys Pro Pro Leu  
 35 40 45  
 Ser Arg Leu Arg Ser Ser His Ala  
 50 55
- 60 (2) INFORMATION FOR SEQ ID NO:31:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 713 base pairs  
 (B) TYPE: nucleic acid
- 612  
 (A) LENGTH: 59 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..59  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011833
- 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:  
 Met Met Val Met Ile Val Ala Val Thr Ile Glu Ala Gln Glu Ser  
 1 5 10 15  
 Gly Trp Thr Ile Cys Phe Arg Gln Cys Ser Gln Pro Cys Arg Ser Asp  
 20 25 30  
 Asp Gly Ser Cys Tyr Glu Asn Cys Lys Ile Glu Cys Gly Gly Pro Lys  
 35 40 45  
 Pro Pro Leu Ser Arg Leu Arg Ser Ser His Ala  
 50 55
- 15 (2) INFORMATION FOR SEQ ID NO:29:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 59 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ix) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..58  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011834
- 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:  
 Met Val Met Ile Val Ala Val Thr Ile Glu Ala Gln Glu Ser Gly  
 1 5 10 15  
 Trp Thr Ile Cys Phe Arg Gln Cys Ser Gln Pro Cys Arg Ser Asp Asp  
 20 25 30  
 Gly Ser Cys Tyr Glu Asn Cys Lys Ile Glu Cys Gly Gly Pro Lys Pro  
 35 40 45  
 Pro Leu Ser Arg Leu Arg Ser Ser His Ala  
 50 55
- 25 (2) INFORMATION FOR SEQ ID NO:30:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 56 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ix) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..56  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011835
- 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:  
 Met Ile Val Ala Val Thr Ile Glu Ala Gln Glu Ser Gly Trp Thr  
 1 5 10 15  
 Ile Cys Phe Arg Gln Cys Ser Gln Pro Cys Arg Ser Asp Gly Ser  
 20 25 30  
 Cys Tyr Glu Asn Cys Lys Ile Glu Cys Gly Pro Lys Pro Pro Leu  
 35 40 45  
 Ser Arg Leu Arg Ser Ser His Ala  
 50 55
- 35 (2) INFORMATION FOR SEQ ID NO:31:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 713 base pairs  
 (B) TYPE: nucleic acid
- 40 (2) INFORMATION FOR SEQ ID NO:30:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 56 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ix) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..56  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011835
- 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:  
 Met Ile Val Ala Val Thr Ile Glu Ala Gln Glu Ser Gly Trp Thr  
 1 5 10 15  
 Ile Cys Phe Arg Gln Cys Ser Gln Pro Cys Arg Ser Asp Gly Ser  
 20 25 30  
 Cys Tyr Glu Asn Cys Lys Ile Glu Cys Gly Pro Lys Pro Pro Leu  
 35 40 45  
 Ser Arg Leu Arg Ser Ser His Ala  
 50 55
- 50 (2) INFORMATION FOR SEQ ID NO:31:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 713 base pairs  
 (B) TYPE: nucleic acid
- 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:  
 Met Ile Val Ala Val Thr Ile Glu Ala Gln Glu Ser Gly Trp Thr  
 1 5 10 15  
 Ile Cys Phe Arg Gln Cys Ser Gln Pro Cys Arg Ser Asp Gly Ser  
 20 25 30  
 Cys Tyr Glu Asn Cys Lys Ile Glu Cys Gly Pro Lys Pro Pro Leu  
 35 40 45  
 Ser Arg Leu Arg Ser Ser His Ala  
 50 55
- 60 (2) INFORMATION FOR SEQ ID NO:28:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 713 base pairs  
 (B) TYPE: nucleic acid

613

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(11) MOLECULE TYPE: DNA (genomic)  
(1x) FEATURE:

5

(A) NAME/KEY: -  
(B) LOCATION: 1..713

(D) OTHER INFORMATION: / Ceres Seq. ID 1011907

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

10 aattctgggt ttcttgcga ttctctccc attctctcc tctctctcc tctctctct 60  
ctctctccg atcatgtccc ctgattccaa atccatacc ttgctctct cttttccaa 120  
ttgattacc gagatttca gatcgagatc tccgattat taataatcc agaaaaaat 180  
aataaaccc gagagagaga gaagatata tatatacatg gagagttcat tagttttcat 240  
tctctccgat tcatatgaca agttcgaatt cgaatacga gcgcagaaaa aatgggtgat 300  
tctctccgat tcatatgaca agttcgaatt cgaatacga gcgcagaaaa aatgggtgat 360  
tgcgcgaga catcttcca aggttaaatc gaattcaag atgcgcgcaa ctatttgact 420  
tgaaggttt ttgtacaaaa tttaattgt aattattca ttgggttt ttgtatttga 480  
aatcgtaaa ctttaattcg gaactgaat ctgggctaaa cttttcaat ctctccacgg 540  
gcaattctt ttctctctc ttcttctct tgaagggtt ttctttgtg ttgtacctg 600  
taataaagc tttgtgtgt gtatacaaa aatgtgattg aataacct ttt 660

(12) INFORMATION FOR SEQ ID NO:32:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 86 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..86

(D) OTHER INFORMATION: / Ceres Seq. ID 1011908

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

35 Met Glu Ser Ser Leu Gly Phe Met Ala Val Phe Ala Val Ser Gly Ser 1  
Val Val Phe Leu Ala Sar Gln Phe His Lys Arg Leu Leu Ser Asp Tyr 15  
Met Asp Lys Phe Glu Phe Glu Ile Arg Ala Gln Lys Lys Met Val Met 20  
Lys Lys Lys Val Arg Phe Ala Ala Asp Val Val Glu Pro Ser Gly Asn 35  
Asn Lys Lys Glu Tyr Arg Arg Arg His Ser Ser Lys Ala Lys Ser Asn Ser 50  
Lys Met Ala Ala Thr Ile 65  
Lys Met Ala Ala Thr Ile 85

(12) INFORMATION FOR SEQ ID NO:33:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 79 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..79

(D) OTHER INFORMATION: / Ceres Seq. ID 1011909

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

55 Met Ala Val Phe Ala Val Ser Gly Ser Val Val Phe Leu Ala Ser Gln 1  
Phe His Lys Arg Leu Leu Ser Asp Tyr Met Asp Lys Phe Glu Phe Glu 10  
Ile Arg Ala Gln Lys Lys Met Val Met Lys Lys Lys Val Arg Phe Ala 25  
35 40 45

614

Ala Asp Val Val Glu Pro Ser Gly Asn Asn Lys Lys Glu Tyr Arg Arg Arg 50  
His Ser Ser Lys Ala Lys Ser Asn Ser Lys Met Ala Ala Thr Ile 60  
65 70 75

5

(2) INFORMATION FOR SEQ ID NO:34:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 59 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..59

(D) OTHER INFORMATION: / Ceres Seq. ID 1011910

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

15 Phe Trp Phe Phe Leu Gln Cys Phe Phe His Phe Ser Pro Pro Leu Ser 1  
Ser Ser Val Ser Leu Leu Arg Ser Cys Pro Leu Ile Pro Asn Pro Ile 5  
Pro Ser Ser Ser Leu Phe Pro Ile Asp Tyr Arg Arg Ile Ser Asp Arg 10  
Asp Leu Arg Phe Ile Asn Asn Tyr Arg Lys Lys 15  
50 55

(12) INFORMATION FOR SEQ ID NO:35:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 580 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: DNA (genomic)

(1x) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..580

(D) OTHER INFORMATION: / Ceres Seq. ID 1011911

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

35 aattgtttta aaattacaaa ttgtccgctt cttttatcc cgtactcgtt cttttttt 60  
ctttcttcc ttatcgcat ttctcgatt ctctcttc cgtccacga ctaatttga 120  
ataaggtta tcaaaagaat aagataagt ggataaag ctactttga aagatttat 180  
gcagagaaa aaaatggat cgagaggat tatcaaat aagtggtcaa tgaggttct 240  
atgggttgtt gctatcgga gtgctatgg ttatacat gttctgtag agagacaaac 300  
tcagaacagg gctcgtgcta tggctgagc ttgagagct cttgaaatc aagtgatgg 360  
tgataatgc taatctac caagtatgc tcaattgat acttcagtt gattttttt 420  
ttttgtgtt tgttttgtt ataagactt cttctgcaa gatgtgttg atgtatttc 480  
tttttcaa ataactgaa taaggtttcg aaacttgag agttgaagt gctgaacata 540  
cgatttgtt tatgcaaaa aagattttt cttatgctg

(12) INFORMATION FOR SEQ ID NO:36:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 59 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(1x) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..59

(D) OTHER INFORMATION: / Ceres Seq. ID 1011912

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

60 Met Gly Ser Arg Gly Ile Ile Asn Asp Lys Tyr Trp Ser Met Arg Ile Leu 1  
Trp Gly Cys Ala Ile Gly Ser Ala Ile Gly Leu Tyr Met Val Ala Val 5  
10 15 20 25 30

615

Glu Arg Gln Thr Gln Asn Arg Ala Arg Ala Met Ala Glu Ser Leu Arg  
35 45  
Ala Ala Glu Ser Gln Gly Asp Gly Asp Asn Val  
50 55

5 (2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 47 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..47

(D) OTHER INFORMATION: / Ceres Seq. ID 1011913

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Met Arg Ile Leu Trp Gly Cys Ala Ile Gly Ser Ala Ile Gly Leu Tyr  
1 10 15

Met Val Ala Val Glu Arg Gln Thr Gln Asn Arg Ala Arg Ala Met Ala  
20 30

Glu Ser Leu Arg Ala Ala Glu Ser Gln Gly Asp Gly Asp Asn Val  
35 40 45

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 40 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..40

(D) OTHER INFORMATION: / Ceres Seq. ID 1011914

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Ile Val Leu Lys Leu Gln Ile Ser Pro Phe Thr Tyr Ser Arg Thr Arg  
1 10 15

Ser Phe Phe Phe Phe Leu Ser Ser Ser Phe Ser Arg Phe Ser Leu  
20 25 30

Phe Arg Ser Pro Thr Asn Ser Glu  
35 40

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 415 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..415

(D) OTHER INFORMATION: / Ceres Seq. ID 1011954

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

tcacaaacc tgattgtct gtttcgaacc gatagattcg aacttggtt taagtgcgcg  
60

gtctcgagga atctccggtt tggacttca cttctgttc ttcttcata cgaagttgt  
120

ttctcttga ataacgttc ttgaacctt ggaatacgtt accagcttc cttcaatcct  
180

taactctagc tcgaagaca tctttgaaa cacttgatc tttaagacc gtttgcaga  
240

cacgagataa gtttgattcg atgtcaacaa cattatttg ccaagggat tgaagcatg  
300

catcctttt agctcaatg attctcrgaa tatgtctcg attcttgga ttgtgatgt  
360

catgatcgc catctgtgc gatattttt ataatgttc gtctgtaagt tattt

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 42 amino acids

616

(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

5

(A) NAME/KEY: peptide

(B) LOCATION: 1..42

(D) OTHER INFORMATION: / Ceres Seq. ID 1011955

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Ser Gln Asn Leu Ile Val Leu Phe Arg Thr Asp Arg Phe Glu Pro Trp  
1 5 10 15

Phe Lys Leu Pro Ala Ser Arg Asn Leu Arg Cys Gly Leu Ser Leu Leu  
20 25 30

Phe Leu Leu His Asp Ala Ser Cys Phe Leu  
35 40

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..32

(D) OTHER INFORMATION: / Ceres Seq. ID 1011956

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Met Thr Gln Ala Val Phe Ser Arg Ile Arg Phe Leu Arg Thr Leu Glu  
1 5 10 15

Asn Val Thr Gln Leu Pro Gln Ser Ser Ser Ser Lys Asn Ile  
20 25 30

Leu

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 440 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..440

(D) OTHER INFORMATION: / Ceres Seq. ID 1011960

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

gacactcagt cactgtaaca tttagatct ttccgaaga agaaacgaa gaagagacga  
60

agagagaaat gagccgatg cagctgatac tttatcaga gatgatgat cgaggtctt  
120

cgatggccat ggaactgat gacctgaag cactggaag actcaacgaa ggagacttg  
180

ttcagataa caagctgcgc gacccgatt tctcaacaa attcgaatt gattcga  
240

acaccgatc caactaaac cggtaagtt ttctcttaa gattcagtt ttttctta  
300

aactcgagt aatttggga gaaccatg tagtaagtg gtaataagt ctgggattt  
360

tatgtttg gtagtaaca atgttttct tcaattggt tactaaaca ttgatttgt  
420

gtgtattgt gttttgtcg

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 62 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

617

(B) LOCATION: 1..62  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011961  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:  
 Met Arg Pro Met Gln Leu Asp Met Leu Ser Glu Met Asp Asp Ala Gly  
 1 5 10 15  
 Ser Ser Met Ala Met Asp Val Asp Asp Leu Glu Ala Met Glu Ile Leu  
 20 25 30  
 Asn Glu Gly Gly Leu Val Ser Asp Asn Lys Leu Ala Asp Ala Asp Phe  
 35 40 45  
 Phe Asn Lys Phe Asp Asp Phe Asp Asp Thr Asp Ile Asn  
 50 55  
 (2) INFORMATION FOR SEQ ID NO:44:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 59 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..59  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011962  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:  
 Met Gln Leu Asp Met Leu Ser Glu Met Asp Asp Ala Gly Ser Ser Met  
 1 5 10 15  
 Ala Met Asp Val Asp Asp Leu Glu Ala Met Glu Ile Leu Asn Glu Gly  
 20 25 30  
 Gly Leu Val Ser Asp Asn Lys Leu Ala Asp Ala Asp Phe Phe Asn Lys  
 35 40 45  
 Phe Asp Asp Phe Asp Asp Thr Asp Ile Asn  
 50 55  
 (2) INFORMATION FOR SEQ ID NO:45:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 55 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..55  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1011963  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:  
 Met Leu Ser Glu Met Asp Asp Ala Gly Ser Ser Met Ala Met Asp Val  
 1 5 10 15  
 Asp Asp Leu Glu Ala Met Glu Ile Leu Asn Glu Gly Leu Val Ser  
 20 25 30  
 Asp Asn Lys Leu Ala Asp Ala Asp Phe Phe Asn Lys Phe Asp Asp Asp  
 35 40 45  
 Phe Asp Asp Thr Asp Ile Asn  
 50 55  
 (2) INFORMATION FOR SEQ ID NO:46:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 193 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..193  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1014075

618

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:  
 aaacccctta agctaaagaa atttcaagc aattcaaaa cctagacgc tctctcttta  
 60  
 tcttcacaa caaatctttt gaattggcga tctctaaagc tctctatttt gtctctatga  
 120  
 tggatattat cctgcgtgtt gcatcgagac agtctgaggc accagcacc aagctctactt  
 180  
 ctgcatctag tcg  
 (2) INFORMATION FOR SEQ ID NO:47:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 39 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..39  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1014076  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:  
 Lys Thr Leu Lys Leu Ser Lys Phe Gln Ser Asn Ser Lys Thr Leu Ala  
 1 5 10 15  
 Ser Leu Ser Leu Ser Leu Thr Thr Lys Ser Leu Lys Trp Arg Ser Leu  
 20 25 30  
 Arg Leu Pro Leu Leu Phe Ser  
 35  
 (2) INFORMATION FOR SEQ ID NO:48:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 36 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..36  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1014077  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:  
 Met Ala Ile Ser Lys Ala Ser Ile Val Val Leu Met Met Val Ile Ile  
 1 5 10 15  
 Ser Val Val Ala Ser Ala Gln Ser Glu Ala Pro Ala Pro Ser Pro Thr  
 20 25 30  
 Ser Gly Ser Ser  
 35  
 (2) INFORMATION FOR SEQ ID NO:49:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 711 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..711  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1015865  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:  
 attcttcgg tcacgcacta attctgttta gctcacattc cttatgtctc ttcattcaca  
 60  
 gtccactaat taacgtacag aatttgttt cctctcttat tctctcttc cttctgaga  
 120  
 ataaagttta tcaaaagaaat aagaataaag tttactcttt aaattccga taattagcag  
 180  
 agtttttca ggattggatt tgatttgcgt ttttggaaac ataatattg cgttttagtg  
 240  
 gataaaagc tagcttaag agttattgca gagaataaaa atcggaatcg gaggtattat  
 300  
 caacgataag tgcctcaatg ggattctatg gggttgtgtg atcggaatg ctattggttt  
 360  
 atacaggttt gctgtagaga gacaaactca gaacagggct cgtgtatgg ctgagagttt  
 420  
 gagaactact gaacacagc gtgatggta taatgtctca tcttcacaa gtatgtctca  
 480  
 gtgtgaactc ctccgttcag tttttttt ttgtgttgt tttgttata argacttctt  
 540

619  
ctgccaagat ggtgtgatg tagttctttt ttgcaataa atcgtaataa ggttcgaaa 600  
ctgagagat tgaagtgtt gaacatacga ttgtgttat cgcaaaaaa gttattctt 660  
atgctgtca tgcatagtt gtaattcga ttittaatgc gttattcag c

(2) INFORMATION FOR SEQ ID NO:50:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 59 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..59  
(D) OTHER INFORMATION: / Ceres Seq. ID 1015866

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:  
Met Gly Ser Arg Gly Ile Ile Asn Asp Lys Trp Ser Met Arg Ile Leu 1  
1 5 10 15  
Trp Gly Cys Ala Ile Gly Ser Ala Ile Gly Leu Tyr Met Val Ala Val 20  
20 25 30  
Glu Arg Gln Thr Gln Asn Arg Ala Arg Ala Met Ala Glu Ser Leu Arg 35  
35 40 45  
Ala Ala Glu Ser Gln Gly Asp Gly Asp Asn Val 50  
50 55

(2) INFORMATION FOR SEQ ID NO:51:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 47 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..47  
(D) OTHER INFORMATION: / Ceres Seq. ID 1015867

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:  
Met Arg Ile Leu Trp Gly Cys Ala Ile Gly Ser Ala Ile Gly Leu Tyr 1  
1 5 10 15  
Met Val Ala Val Glu Arg Gln Thr Gln Asn Arg Ala Arg Ala Met Ala 20  
20 25 30  
Glu Ser Leu Arg Ala Ala Glu Ser Gln Gly Asp Gly Asp Asn Val 35  
35 40 45

(2) INFORMATION FOR SEQ ID NO:52:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 38 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..38  
(D) OTHER INFORMATION: / Ceres Seq. ID 1015868

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:  
Ser Ser Gly His Arg Leu Ile Leu Val Ser Ser His Ser Leu Cys Ser 1  
1 5 10 15  
Phe Ile His Ser Ser Leu Ile Asn Ala Thr Asn Phe Val Ser Pro Leu 20  
20 25 30  
Ile Ser Ser Phe Leu Leu 35

(2) INFORMATION FOR SEQ ID NO:53:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 432 base pairs

620  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..432  
(D) OTHER INFORMATION: / Ceres Seq. ID 1021371

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:  
gtcattttt cgattctac tcttcggtc accgactaat tctgagaaa aaaatggat 60  
60 120  
cgaggggat tctcaacgat aagtggtcaa tgggattctt acgggttgtt gctatcgaa 120  
120 180  
gtgtattgg ttatcacatg gtgtgtgag adagacaaac tcaaacagc gcctgcgca 180  
180 240  
tgcctggag ttgagagct gctgaatcac aagtgatgg tgataatgta taattctac 240  
240 300  
caagtagtc tcaattgaat actctcagtt gattttttt tttagtggtt tgttttgtt 300  
300 360  
ataatgact ctctcgcca gatgggttg atgtagtttc tttttgcaa ataactgtaa 360  
360 420  
taagtgttc aaactggag agttgaagt gctgaacata cgatttggt ratcgcaaaa 420  
aaagtattt cc

(2) INFORMATION FOR SEQ ID NO:54:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 75 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..75  
(D) OTHER INFORMATION: / Ceres Seq. ID 1021372

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:  
His Phe Leu Asp Ser His Ser Gly His Arg Leu Ile Leu Arg Lys 1  
1 5 10 15  
Lys Met Gly Ser Arg Gly Ile Ile Asn Asp Lys Trp Ser Met Arg Ile 20  
20 25 30  
Leu Trp Gly Cys Ala Ile Gly Ser Ala Ile Gly Leu Tyr Met Val Ala 35  
35 40 45  
Val Glu Arg Gln Thr Gln Asn Arg Ala Arg Ala Met Ala Glu Ser Leu 50  
50 55 60  
Arg Ala Ala Glu Ser Gln Gly Asp Gly Asp Asn Val 65  
65 70 75

(2) INFORMATION FOR SEQ ID NO:55:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 59 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..59  
(D) OTHER INFORMATION: / Ceres Seq. ID 1021373

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:  
Met Gly Ser Arg Gly Ile Ile Asn Asp Lys Trp Ser Met Arg Ile Leu 1  
1 5 10 15  
Trp Gly Cys Ala Ile Gly Ser Ala Ile Gly Leu Tyr Met Val Ala Val 20  
20 25 30  
Glu Arg Gln Thr Gln Asn Arg Ala Arg Ala Met Ala Glu Ser Leu Arg 35  
35 40 45  
Ala Ala Glu Ser Gln Gly Asp Gly Asp Asn Val 50  
50 55

(2) INFORMATION FOR SEQ ID NO:56:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 47 amino acids

622

(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(11) MOLECULE TYPE: peptide  
(1x) FEATURE:

(B) LOCATION: 1...47  
 (D) OTHER INFORMATION: / Ceres Seq. ID I021374  
 (x1) SEQUENCE DESCRIPTION: SEQ ID NO:56:  
 Met Arg Ile Leu Trp Gly Cys Ala Ile Gly Ser Ala Ile Gly Leu Tyr  
 1 5 10 15  
 Met Val Ala Val Glu Arg Gln Thr Gln Asn Arg Ala Arg Ala Met Ala  
 20 25 30

15

(2) INFORMATION FOR SEQ ID NO:57:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 286 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (i) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:

25

(B) LOCATION: 1..285  
(D) OTHER INFORMATION: / Ceres Seq. ID 1022578  
(x1) SEQUENCE DESCRIPTION: SEQ ID NO:57:

60           accatcgttga tgtctcggt atacacgtgt aaacttccca ggaacaggat tctgtggtaaa  
 120           gcaaaagaac tgaatcga acaaaactct ggtttggtt caaaatcagc gaagaagaac  
 180           tctgttggtt ggtttggtt gaatttcaa cyttgatcgt tgaattatat gtcaactctt  
 240           agaatccga aagtttgcgc attatcagga aagtttgcgc agacgcagtc aggtttcgtg tctgaggattc  
 300           atcaaaaatt tgttaacttt attcgaaatt catgtttatt cggcga

35

(A) LENGTH: 43 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(E) MOLECULE TYPE: peptide  
(F) FEATURE:

(A) NAME/KEY: peptide  
(B) LOCATION: 1..43

(D) OTHER INFORMATION: / Ceres Seq. ID 1022579

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

His	His	Trp	Met	Val	Arg	Leu	Tyr	Lys	Leu	Lys	Pro	Arg	Asn	Arg
1			5										15	
Ile	Arg	Gly	Lys	Ala	Lys	Lys	Leu	Lys	Ser	Lys	Thr	Ser	Ser	Gly
			20						25				30	
Ser	Ser	Lys	Ser	Ala	Lys	Lys	Asn	Pro	Trp	Val				

50

INFORMATION FOR SEQ ID NO:59:  
(1) SEQUENCE CHARACTERISTICS  
    (A) LENGTH: 43 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS:

(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:

60

- (A) NAME/KEY: peptide
- (B) LOCATION: 1..43
- (D) OTHER INFORMATION: / Ceres Seq. ID 1022580
- (x1) SEQUENCE DESCRIPTION: SEQ ID NO:59:

16C:ON AT 09Z

Met Val Ser Tyr Met Ser Thr Leu Arg Ser Leu Tyr Tyr Arg Glu Ser  
1 5 10 15  
Cys Cys Arg Arg Ser Gln Val Ser Gly Val Gly Phe Val Lys Ile Leu  
20 25 30  
5 Leu Thr Leu Phe Glu Phe His Val Tyr Cys Gly

```

(2) INFORMATION FOR SEQ ID NO:60:
      (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 40 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS:
            (D) TOPOLOGY: linear
      (i) MOLECULE TYPE: peptide

```

15

```

(A) NAME/KEY: peptide
(B) LOCATION: 1..40
(D) OTHER INFORMATION: / Ceres Seq. ID: I027581
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

```

20 Met Val Arg Leu Tyr Lys Leu Lys Pro Arg Asn Arg Ile Arg Gly  
1  
Lys Ala Lys Lys Leu Lys Ser Lys Thr Ser Ser Gly Leu Ser Ser Lys  
20 25 30

25

```
(i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 643 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(ix) FEATURE:
```

35

```
(D) OTHER INFORMATION: / Ceres Seq. ID 1024240
(x1) SEQUENCE DESCRIPTION: SEQ ID NO:61:
tcacgg agattctt aggcagattt ctctgaaa tcgtttcagg aaaaatgaga 60
tatcag gataccag aatctcaagt catcgaagt taccgcana gagaagatc 120
gaatc aaaaacccat tcacagct atcagaggt atctatctga agtcaggcic 180
tactt ctgtctgat attagacat ttgtctcga gctctcggg taagtattga 240
gataag cgcacaaagc tctcgttt ttgacgcga aaaaagaagt tccctctcag 300
gataag cgtgaacct taaatatta agggcgtggt gatctatct agggattaaa 360
agauc ccaaaaggg cagcagaatt tgcctgnaac ataagattt tcaggaga 420
gatga cggatcaaa ttgtctcaga atgtttgat aagtcctct atcagcagt 480
gsaac atttgatct tcacgcgna catcnaagg aataagacct tgcgtatga 540
taactg carttttc aatcatcct tgguaaagt aatgaacca cgcgtatga 600
atttta attccacca ttaatcaat tcagagact tcc
INFORMATION FOR SEQ ID NO:62:
```

50

- (ii) MOLECULE TYPE: peptide
- (ix) FEATURE:

(A) NAME/KEY: peptide  
(B) LOCATION: 1..59  
(D) OTHER INFORMATION: / Ceres Seq. ID 1024241  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:  
Glu Thr Ser Met Arg Tyr Thr Ser Asn Ser Lys Ser Met Lys Ile  
5 10 15  
Ala Lys Glu Lys Val Pro Val Asn Ser Lys Thr His Leu Glu Leu



623

His Gly Glu Leu Asp Thr Gly Thr Gly Ala Pro Ser Tyr Phe Cys Glu 30  
 35  
 Met Ile Arg His Phe Phe Pro Glu Ala Leu Gly 45  
 50  
 5

## (2) INFORMATION FOR SEQ ID NO:63:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 55 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..55

(D) OTHER INFORMATION: / Ceres Seq. ID 1024242

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

Met Arg Tyr Thr Ser Asn Ser Lys Ser Met Lys Ile His Ala Lys Glu 10  
 15  
 Lys Val Pro Val Asn Ser Lys Thr His Leu Gln Leu His Gly Glu Leu 20  
 25  
 Asp Thr Gly Thr Gly Ala Pro Ser Tyr Phe Cys Ala Met Ile Arg His 35  
 45  
 Phe Phe Pro Glu Ala Leu Gly 40  
 50  
 25

## (2) INFORMATION FOR SEQ ID NO:64:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 47 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..47

(D) OTHER INFORMATION: / Ceres Seq. ID 1024243

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

Met Asp Phe Lys Glu Asp Gln Asp Val Arg Ile Lys Val Gly Tyr Glu 15  
 20  
 Met Phe Asp Lys Val Pro Tyr Met Gln Ile Arg Glu Asn Asn Trp Thr 25  
 30  
 Leu Asn Ala Asn Met Lys Gly Lys Trp Asn Leu Arg Tyr Asp Leu 35  
 40  
 45

## (2) INFORMATION FOR SEQ ID NO:65:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 729 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..729

(D) OTHER INFORMATION: / Ceres Seq. ID 1026562

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

gtaacata taattgac cctaaatct atctatct ctcataatc tttctctt 60  
 aaatccctaa atcttcagt tactttcta tccacattt cacttcaat ggccttatt 120  
 atcaactgt ccgctctacc cagctacga gcatcttcg gaccggatc ccttaaccg 180  
 gatacaaac gaagaatc tgcgtcttg tggcttcct tcttcgttt acctccgat 240  
 ccagattacc tcaacatga aagctcatgc tccaccgtga atcggataa aaccgatatt 300  
 tccggatcgg gtcaaaagt tccgtccggg tctttaccg aggaagaagc taagcagtt 360  
 aggaagaaac ccgcagaagc ttcacagtc catgaogtaa tgatcaactc cgcactgtc 420

624

ttcgacttg cgtctgatat cacggccga gtcgagatt gaacggatt cgggtcagat 480  
 ctggagaga acaattaat gaataagaa ttaattttt ttctctaat tgaattgtct 540  
 tggagagc agacatctt ctatcgctgc tctctttca ttggaacgt tggattgtgt 600  
 ttgctcttg tgtgtgtaa acctgtgtc cctttctgt tctctgtaa tcaaccttgg 660  
 gatttccact tgtattttt agacatgtgt tccctagtagt gctctttat ctatactaa 720  
 atttaagc

## (2) INFORMATION FOR SEQ ID NO:66:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 153 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..153

(D) OTHER INFORMATION: / Ceres Seq. ID 1026563

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

Val Asn His Ile Asn Leu Thr Leu Lys Ser Ile Ser Phe Ser His Gln 10  
 15  
 Ser Ser Ser Phe Lys Ser Leu Asn Leu His Val Thr Phe Leu Ser Thr 20  
 25  
 Phe Ser Leu Ser Met Ala Leu Ile Ile Thr Cys Ser Ala Leu Pro Thr 30  
 35  
 Ile Arg Ala Ser Ser Gly Ser Gly Ser Leu Asn Pro Asp Gln Asn Arg 40  
 45  
 Lys Lys Ser Ala Ala Trp Trp Ala Pro Leu Phe Gly Leu Pro Ser Asp 50  
 55  
 Pro Asp Tyr Leu Asn Ile Glu Ser Ser Cys Ser Thr Val Asn Pro Asp 60  
 65  
 Lys Thr Asp Ile Ser Gly Ser Gly Gln Lys Phe Arg Arg Gly Cys Phe 70  
 75  
 Thr Glu Glu Lys Ala Lys Gln Leu Phe Arg Lys Thr Ala Glu Ala Ser 80  
 85  
 Thr Phe His Asp Val Met Tyr His Ser Ala Ile Ala Ser Arg Leu Ala 90  
 95  
 Ser Asp Ile Thr Gly Arg Val Glu Asp 100  
 105  
 110  
 115  
 120  
 125  
 130  
 135  
 140  
 145  
 150  
 155  
 160  
 165  
 170  
 175  
 180  
 185  
 190  
 195  
 200  
 205  
 210  
 215  
 220  
 225  
 230  
 235  
 240  
 245  
 250  
 255  
 260  
 265  
 270  
 275  
 280  
 285  
 290  
 295  
 300  
 305  
 310  
 315  
 320  
 325  
 330  
 335  
 340  
 345  
 350  
 355  
 360  
 365  
 370  
 375  
 380  
 385  
 390  
 395  
 400  
 405  
 410  
 415  
 420  
 425  
 430  
 435  
 440  
 445  
 450  
 455  
 460  
 465  
 470  
 475  
 480  
 485  
 490  
 495  
 500  
 505  
 510  
 515  
 520  
 525  
 530  
 535  
 540  
 545  
 550  
 555  
 560  
 565  
 570  
 575  
 580  
 585  
 590  
 595  
 600  
 605  
 610  
 615  
 620  
 625  
 630  
 635  
 640  
 645  
 650  
 655  
 660  
 665  
 670  
 675  
 680  
 685  
 690  
 695  
 700  
 705  
 710  
 715  
 720  
 725  
 730  
 735  
 740  
 745  
 750  
 755  
 760  
 765  
 770  
 775  
 780  
 785  
 790  
 795  
 800  
 805  
 810  
 815  
 820  
 825  
 830  
 835  
 840  
 845  
 850  
 855  
 860  
 865  
 870  
 875  
 880  
 885  
 890  
 895  
 900  
 905  
 910  
 915  
 920  
 925  
 930  
 935  
 940  
 945  
 950  
 955  
 960  
 965  
 970  
 975  
 980  
 985  
 990  
 995  
 1000  
 1005  
 1010  
 1015  
 1020  
 1025  
 1030  
 1035  
 1040  
 1045  
 1050  
 1055  
 1060  
 1065  
 1070  
 1075  
 1080  
 1085  
 1090  
 1095  
 1100  
 1105  
 1110  
 1115  
 1120  
 1125  
 1130  
 1135  
 1140  
 1145  
 1150  
 1155  
 1160  
 1165  
 1170  
 1175  
 1180  
 1185  
 1190  
 1195  
 1200  
 1205  
 1210  
 1215  
 1220  
 1225  
 1230  
 1235  
 1240  
 1245  
 1250  
 1255  
 1260  
 1265  
 1270  
 1275  
 1280  
 1285  
 1290  
 1295  
 1300  
 1305  
 1310  
 1315  
 1320  
 1325  
 1330  
 1335  
 1340  
 1345  
 1350  
 1355  
 1360  
 1365  
 1370  
 1375  
 1380  
 1385  
 1390  
 1395  
 1400  
 1405  
 1410  
 1415  
 1420  
 1425  
 1430  
 1435  
 1440  
 1445  
 1450  
 1455  
 1460  
 1465  
 1470  
 1475  
 1480  
 1485  
 1490  
 1495  
 1500  
 1505  
 1510  
 1515  
 1520  
 1525  
 1530  
 1535  
 1540  
 1545  
 1550  
 1555  
 1560  
 1565  
 1570  
 1575  
 1580  
 1585  
 1590  
 1595  
 1600  
 1605  
 1610  
 1615  
 1620  
 1625  
 1630  
 1635  
 1640  
 1645  
 1650  
 1655  
 1660  
 1665  
 1670  
 1675  
 1680  
 1685  
 1690  
 1695  
 1700  
 1705  
 1710  
 1715  
 1720  
 1725  
 1730  
 1735  
 1740  
 1745  
 1750  
 1755  
 1760  
 1765  
 1770  
 1775  
 1780  
 1785  
 1790  
 1795  
 1800  
 1805  
 1810  
 1815  
 1820  
 1825  
 1830  
 1835  
 1840  
 1845  
 1850  
 1855  
 1860  
 1865  
 1870  
 1875  
 1880  
 1885  
 1890  
 1895  
 1900  
 1905  
 1910  
 1915  
 1920  
 1925  
 1930  
 1935  
 1940  
 1945  
 1950  
 1955  
 1960  
 1965  
 1970  
 1975  
 1980  
 1985  
 1990  
 1995  
 2000  
 2005  
 2010  
 2015  
 2020  
 2025  
 2030  
 2035  
 2040  
 2045  
 2050  
 2055  
 2060  
 2065  
 2070  
 2075  
 2080  
 2085  
 2090  
 2095  
 2100  
 2105  
 2110  
 2115  
 2120  
 2125  
 2130  
 2135  
 2140  
 2145  
 2150  
 2155  
 2160  
 2165  
 2170  
 2175  
 2180  
 2185  
 2190  
 2195  
 2200  
 2205  
 2210  
 2215  
 2220  
 2225  
 2230  
 2235  
 2240  
 2245  
 2250  
 2255  
 2260  
 2265  
 2270  
 2275  
 2280  
 2285  
 2290  
 2295  
 2300  
 2305  
 2310  
 2315  
 2320  
 2325  
 2330  
 2335  
 2340  
 2345  
 2350  
 2355  
 2360  
 2365  
 2370  
 2375  
 2380  
 2385  
 2390  
 2395  
 2400  
 2405  
 2410  
 2415  
 2420  
 2425  
 2430  
 2435  
 2440  
 2445  
 2450  
 2455  
 2460  
 2465  
 2470  
 2475  
 2480  
 2485  
 2490  
 2495  
 2500  
 2505  
 2510  
 2515  
 2520  
 2525  
 2530  
 2535  
 2540  
 2545  
 2550  
 2555  
 2560  
 2565  
 2570  
 2575  
 2580  
 2585  
 2590  
 2595  
 2600  
 2605  
 2610  
 2615  
 2620  
 2625  
 2630  
 2635  
 2640  
 2645  
 2650  
 2655  
 2660  
 2665  
 2670  
 2675  
 2680  
 2685  
 2690  
 2695  
 2700  
 2705  
 2710  
 2715  
 2720  
 2725  
 2730  
 2735  
 2740  
 2745  
 2750  
 2755  
 2760  
 2765  
 2770  
 2775  
 2780  
 2785  
 2790  
 2795  
 2800  
 2805  
 2810  
 2815  
 2820  
 2825  
 2830  
 2835  
 2840  
 2845  
 2850  
 2855  
 2860  
 2865  
 2870  
 2875  
 2880  
 2885  
 2890  
 2895  
 2900  
 2905  
 2910  
 2915  
 2920  
 2925  
 2930  
 2935  
 2940  
 2945  
 2950  
 2955  
 2960  
 2965  
 2970  
 2975  
 2980  
 2985  
 2990  
 2995  
 3000  
 3005  
 3010  
 3015  
 3020  
 3025  
 3030  
 3035  
 3040  
 3045  
 3050  
 3055  
 3060  
 3065  
 3070  
 3075  
 3080  
 3085  
 3090  
 3095  
 3100  
 3105  
 3110  
 3115  
 3120  
 3125  
 3130  
 3135  
 3140  
 3145  
 3150  
 3155  
 3160  
 3165  
 3170  
 3175  
 3180  
 3185  
 3190  
 3195  
 3200  
 3205  
 3210  
 3215  
 3220  
 3225  
 3230  
 3235  
 3240  
 3245  
 3250  
 3255  
 3260  
 3265  
 3270  
 3275  
 3280  
 3285  
 3290  
 3295  
 3300  
 3305  
 3310  
 3315  
 3320  
 3325  
 3330  
 3335  
 3340  
 3345  
 3350  
 3355  
 3360  
 3365  
 3370  
 3375  
 3380  
 3385  
 3390  
 3395  
 3400  
 3405  
 3410  
 3415  
 3420  
 3425  
 3430  
 3435  
 3440  
 3445  
 3450  
 3455  
 3460  
 3465  
 3470  
 3475  
 3480  
 3485  
 3490  
 3495  
 3500  
 3505  
 3510  
 3515  
 3520  
 3525  
 3530  
 3535  
 3540  
 3545  
 3550  
 3555  
 3560  
 3565  
 3570  
 3575  
 3580  
 3585  
 3590  
 3595  
 3600  
 3605  
 3610  
 3615  
 3620  
 3625  
 3630  
 3635  
 3640  
 3645  
 3650  
 3655  
 3660  
 3665  
 3670  
 3675  
 3680  
 3685  
 3690  
 3695  
 3700  
 3705  
 3710  
 3715  
 3720  
 3725  
 3730  
 3735  
 3740  
 3745  
 3750  
 3755  
 3760  
 3765  
 3770  
 3775  
 3780  
 3785  
 3790  
 3795  
 3800  
 3805  
 3810  
 3815  
 3820  
 3825  
 3830  
 3835  
 3840  
 3845  
 3850  
 3855  
 3860  
 3865  
 3870  
 3875  
 3880  
 3885  
 3890  
 3895  
 3900  
 3905  
 3910  
 3915  
 3920  
 3925  
 3930  
 3935  
 3940  
 3945  
 3950  
 3955  
 3960  
 3965  
 3970  
 3975  
 3980  
 3985  
 3990  
 3995  
 4000  
 4005  
 4010  
 4015  
 4020  
 4025  
 4030  
 4035  
 4040  
 4045  
 4050  
 4055  
 4060  
 4065  
 4070  
 4075  
 4080  
 4085  
 4090  
 4095  
 4100  
 4105  
 4110  
 4115  
 4120  
 4125  
 4130  
 4135  
 4140  
 4145  
 4150  
 4155  
 4160  
 4165  
 4170  
 4175  
 4180  
 4185  
 4190  
 4195  
 4200  
 4205  
 4210  
 4215  
 4220  
 4225  
 4230  
 4235  
 4240  
 4245  
 4250  
 4255  
 4260  
 4265  
 4270  
 4275  
 4280  
 4285  
 4290  
 4295  
 4300  
 4305  
 4310  
 4315  
 4320  
 4325  
 4330  
 4335  
 4340  
 4345  
 4350  
 4355  
 4360  
 4365  
 4370  
 4375  
 4380  
 4385  
 4390  
 4395  
 4400  
 4405  
 4410  
 4415  
 4420  
 4425  
 4430  
 4435  
 4440  
 4445  
 4450  
 4455  
 4460  
 4465  
 4470  
 4475  
 4480  
 4485  
 4490  
 4495  
 4500  
 4505  
 4510  
 4515  
 4520  
 4525  
 4530  
 4535  
 4540  
 4545  
 4550  
 4555  
 4560  
 4565  
 4570  
 4575  
 4580  
 4585  
 4590  
 4595  
 4600  
 4605  
 4610  
 4615  
 4620  
 4625  
 4630  
 4635  
 4640  
 4645  
 4650  
 4655  
 4660  
 4665  
 4670  
 4675  
 4680  
 4685  
 4690  
 4695  
 4700  
 4705  
 4710  
 4715  
 4720  
 4725  
 4730  
 4735  
 4740  
 4745  
 4750  
 4755  
 4760  
 4765  
 4770  
 4775  
 4780  
 4785  
 4790  
 4795  
 4800  
 4805  
 4810  
 4815  
 4820  
 4825  
 4830  
 4835  
 4840  
 4845  
 4850  
 4855  
 4860  
 4865  
 4870  
 4875  
 4880  
 4885  
 4890  
 4895  
 4900  
 4905  
 4910  
 4915  
 4920  
 4925  
 4930  
 4935  
 4940  
 4945  
 4950  
 4955  
 4960  
 4965  
 4970  
 4975  
 4980  
 4985  
 4990  
 4995  
 5000  
 5005  
 5010  
 5015  
 5020  
 5025  
 5030  
 5035  
 5040  
 5045  
 5050  
 5055  
 5060  
 5065  
 5070  
 5075  
 5080  
 5085  
 5090  
 5095  
 5100  
 5105  
 5110  
 5115  
 5120  
 5125  
 5130  
 5135  
 5140  
 5145  
 5150  
 5155  
 5160  
 5165  
 5170  
 5175  
 5180  
 5185  
 5190  
 5195  
 5200  
 5205  
 5210  
 5215  
 5220  
 5225  
 5230  
 5235  
 5240  
 5245  
 5250  
 5255  
 5260  
 5265  
 5270  
 5275  
 5280  
 5285  
 5290  
 5295  
 5300  
 5305  
 5310  
 5315  
 5320  
 5325  
 5330  
 5335  
 5340  
 5345  
 5350  
 5355  
 5360  
 5365  
 5370  
 5375  
 5380  
 5385  
 5390  
 5395  
 5400  
 5405  
 5410  
 5415  
 5420  
 5425  
 5430  
 5435  
 5440  
 5445  
 5450  
 5455  
 5460  
 5465  
 5470  
 5475  
 5480  
 5485  
 5490  
 5495  
 5500  
 55

625

Val Met Tyr His Ser Ala Ile Ala Ser Arg Leu Ala Ser Asp Ile Thr  
100 105  
Gly Arg Val Glu Asp  
115

5 (2) INFORMATION FOR SEQ ID NO:68:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 484 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..484  
(D) OTHER INFORMATION: / Ceres Seq. ID 1026648

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:  
aaggatcaca aaattagggt ttaattga tagagaagt gaatttcaga agctttgagg  
agttcggcc ttctacacg atgcacact cgaatcacg gaaggagaag tggacttca  
taggtatct cggagacatc gttgcttga tatgttcgt ttgatcaac tgggtttct  
20 tagcttggg gctctgctt ggtacacggt tgggtggtt tagcacttc ttcctggaag  
ggaattgtc ggcagcttt ggcacacgc ttggctggt tccctgcgt cccaagatgt  
ttagtctgt gccacagga agcacgaga gagagatgaa gagacttgg agaggccat  
rgttagact ctttgaagc agcagcagat acctctatg aattgtctt gattcttct  
attgcgtct ggaattggt ttgaactcc agtttgtat tacttaccg aatcgttgtt  
tgcg 480

25 (2) INFORMATION FOR SEQ ID NO:69:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 112 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..112  
(D) OTHER INFORMATION: / Ceres Seq. ID 1026649

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:  
Met Asn Phe Arg Ser Phe Glu Phe Thr Pro Phe Tyr Met Met Gln  
1 10 15

40 His Ser Asn Pro Ser Thr Arg Arg Leu His Phe Ile Gly Ile Ala  
20 25 30

Ser Ile Val Ala Leu Ile Cys Ser Ile Leu Ile Asn Trp Trp Phe Leu  
35 40 45

Ala Leu Val Pro Leu Leu Gly Tyr Gly Phe Ala Trp Tyr Ser His Phe  
50 55 60

Phe Val Glu Gly Asn Val Pro Ala Ser Phe Gly His Pro Leu Trp Ser  
65 70 75 80

Phe Leu Cys Asp Leu Lys Met Phe Ser Leu Met Leu Thr Gly Ser Met  
85 90 95

Glu Arg Glu Met Lys Arg Leu Gly Lys Arg Pro Leu Leu Gln Leu Ser  
100 105 110

(2) INFORMATION FOR SEQ ID NO:70:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 99 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..99  
(D) OTHER INFORMATION: / Ceres Seq. ID 1026650

626

Met Met Gln His Ser Asn Pro Ser Thr Arg Arg Leu His Phe Ile Gly  
1 5 10 15  
Ile Ile Ala Ser Ile Val Ala Leu Ile Cys Ser Ile Leu Ile Asn Trp  
20 25 30

5 Trp Phe Leu Ala Leu Val Pro Leu Leu Gly Tyr Phe Ala Trp Tyr  
35 40 45  
Ser His Phe Phe Val Glu Gly Asn Val Pro Ala Ser Phe Gly His Pro  
50 55 60

10 Leu Trp Ser Phe Leu Cys Asp Leu Lys Met Phe Ser Leu Met Leu Thr  
65 70 75 80

Gly Ser Met Glu Arg Glu Met Lys Arg Leu Gly Lys Arg Pro Leu Leu  
85 90 95

Gln Leu Ser  
15

(2) INFORMATION FOR SEQ ID NO:71:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 98 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..98  
(D) OTHER INFORMATION: / Ceres Seq. ID 1026651

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:  
Met Gln His Ser Asn Pro Ser Thr Arg Arg Leu His Phe Ile Gly Ile  
1 5 10 15

30 Ile Ala Ser Ile Val Ala Leu Ile Cys Ser Ile Leu Ile Asn Trp Trp  
20 25 30

Phe Leu Ala Leu Val Pro Leu Leu Gly Tyr Gly Phe Ala Trp Tyr Ser  
35 40 45

His Phe Phe Val Glu Gly Asn Val Pro Ala Ser Phe Gly His Pro Leu  
50 55 60

Trp Ser Phe Leu Cys Asp Leu Lys Met Phe Ser Leu Met Leu Thr Gly  
65 70 75 80

Ser Met Glu Arg Glu Met Lys Arg Leu Gly Lys Arg Pro Leu Leu Gln  
85 90 95

Leu Ser  
40

(2) INFORMATION FOR SEQ ID NO:72:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 563 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..563  
(D) OTHER INFORMATION: / Ceres Seq. ID 1027881

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:  
aaacacacaga gaggtagcag agagagagag tgaatatga gaagtaactt ggaatgcgt  
60 acagggggga tccaggatg ccaatgcag atgcagatcg ttcatgaat atagtgatg  
120 gttcgtgc ttctccgtt ttactcgtg ttaactcta catggcgag cttctaac  
180 agttcaatta ccatgacaag tggatgctg ttgagcagta caactggaa aaagaagg  
240 caaagaagca acctatgaa tccaatgga ataagatcc caagaagtc aggaactcgt  
300 actatacaa ctgcccagtc tacttccat agaagtgtct cgtgttgcgt gtaatcaga  
360 agagaacaa aacacttgc gaactctat ggaatctat tctactctt ttgtatcgt  
420 tatctgagc accttggat tttaagttt ttctttgta atgacaatc tctcaacct  
480 tcaacttct atctcctta gtgactctc ttatggaaat atgaaatctc ttatgaagac  
540

627

agaaccagaa tctgtctgtt atg  
 (2) INFORMATION FOR SEQ ID NO:73:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 109 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..109  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1027882  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:  
 Asn Thr Glu Ser Asp Glu Arg Glu Val Lys Met Glu Lys Tyr Phe  
 1 5 10 15  
 Gly Asn Ala Tyr Arg Gly Asp Pro Gly Val Pro His Ala Asp Ala Asp  
 20 25  
 Arg Phe Met Asn Ile Trp Ile Gly Ser Ala Ala Phe Ser Val Leu Thr  
 30 35 40 45  
 Trp Val Asn Pro Tyr Met Trp Glu Lys Ser Asn Glu Phe Asn Tyr His  
 50 55  
 Asp Lys Trp Met Leu Phe Glu Lys Tyr His Trp Lys Ala Arg Ala  
 60 65 70 75 80  
 Lys Lys Glu Pro Tyr Glu Phe Lys Trp Asn Lys Ile Pro Lys Glu Val  
 85 90 95  
 Arg Asp Ser Tyr Tyr Asn Trp Pro Val Tyr Phe Pro  
 100 105  
 (2) INFORMATION FOR SEQ ID NO:74:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 98 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..98  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1027883  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:  
 Met Glu Lys Tyr Phe Gly Asn Ala Tyr Arg Gly Asp Pro Gly Val Pro  
 1 5 10 15  
 His Ala Asp Ala Asp Arg Phe Met Asn Ile Trp Ile Gly Ser Ala Ala  
 20 25 30  
 Phe Ser Val Leu Thr Trp Val Asn Pro Tyr Trp Met Trp Glu Lys Ser Asn  
 35 40 45  
 Glu Phe Asn Tyr His Asp Lys Trp Met Leu Phe Glu Lys Tyr His Trp  
 50 55 60  
 Lys Lys Ala Arg Ala Lys Lys Glu Pro Tyr Glu Phe Lys Trp Asn Lys  
 65 70 75 80  
 Ile Pro Lys Glu Val Arg Asp Ser Tyr Tyr Asn Trp Pro Val Tyr  
 85 90 95  
 Phe Pro  
 (2) INFORMATION FOR SEQ ID NO:75:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 75 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide

628

(B) LOCATION: 1..75  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1027884  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:  
 Met Asn Ile Trp Ile Gly Ser Ala Ala Phe Ser Val Leu Thr Trp Val  
 1 5 10 15  
 Asn Pro Tyr Met Trp Glu Lys Ser Asn Gln Phe Asn Tyr His Asp Lys  
 20 25 30  
 Trp Met Leu Phe Glu Lys Tyr His Trp Lys Lys Ala Arg Ala Lys Lys  
 35 40 45  
 Gln Pro Tyr Glu Phe Lys Trp Asn Lys Ile Pro Lys Glu Val Arg Asp  
 50 55 60  
 Ser Tyr Tyr Tyr Asn Trp Pro Val Tyr Phe Pro  
 65 70 75  
 (2) INFORMATION FOR SEQ ID NO:76:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 305 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..305  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1381797  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:  
 gaggtttgg ttgttgac tcaatgagga ttcttaggc gattttctcg agaaatcgt  
 60  
 ttgagaaaa tggagacttc tatgagtat accagcaatt ccaagctcat gaagatcat  
 120  
 gccaaaga aggttcgggt gaactcaaaa acccaattac agttcctgg aggtttgat  
 180  
 actggaatg ggggtccggg ttacttcgt gcgatgatta gacactttt tcttgagct  
 240  
 tcaacagcc ttgggttagg attgcattat gataagcgc aaagcttcg gtgtcttgta  
 300  
 cgcgg  
 (2) INFORMATION FOR SEQ ID NO:77:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 78 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..78  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1381798  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:  
 Met Glu Thr Ser Met Arg Tyr Thr Ser Asn Ser Lys Ser Met Lys Ile  
 1 5 10 15  
 His Ala Lys Glu Lys Val Pro Val Asn Ser Lys Thr His Leu Gln Leu  
 20 25 30  
 His Gly Glu Leu Asp Thr Gly Thr Gly Ala Pro Ser Tyr Phe Cys Ala  
 35 40 45  
 Met Ile Arg His Phe Pro Glu Ala Ser Thr Gly Leu Gly Val Gly  
 50 55 60  
 Leu His Tyr Asp Lys Arg Gln Lys Leu Arg Cys Leu Val Arg  
 65 70 75  
 (2) INFORMATION FOR SEQ ID NO:78:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 74 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide

629

(B) LOCATION: 1..74

(D) OTHER INFORMATION: / Ceres Seq. ID 1381799

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

Met Arg Tyr Thr Ser Asn Ser Lys Ser Met Lys Ile His Ala Lys Glu  
 1 10  
 Lys Val Pro Val Asn Ser Lys Thr His Leu Gln Leu His Gly Glu Leu  
 20 30  
 Asp Thr Gly Thr Gly Ala Pro Ser Tyr Phe Cys Ala Met Ile Arg His  
 35 40  
 Phe Phe Pro Glu Ala Ser Thr Gly Leu Gly Val Gly Leu His Tyr Asp  
 50 55  
 Lys Arg Gln Lys Leu Arg Cys Leu Val Arg  
 65 70

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 65 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..65

(D) OTHER INFORMATION: / Ceres Seq. ID 1381800

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

Met Lys Ile His Ala Lys Glu Lys Val Pro Val Asn Ser Lys Thr His  
 1 10  
 Leu Gln Leu His Gly Glu Leu Asp Thr Gly Thr Gly Ala Pro Ser Tyr  
 20 25  
 Phe Cys Ala Met Ile Arg His Phe Phe Pro Glu Ala Ser Thr Gly Leu  
 35 40  
 Gly Val Gly Leu His Tyr Asp Lys Arg Gln Lys Leu Arg Cys Leu Val  
 50 55  
 Arg  
 65

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 840 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..840

(D) OTHER INFORMATION: / Ceres Seq. ID 1442747

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

acacgaacg attccgctc tccgtctct tcataaat ctaactcgc gtcaactac  
 120  
 tcttcttca cagcaaat tcaattctt ctaaccgtc aacgaatcc tcttctcaa  
 180  
 gttttattt cctctcgga tggctcaga agatgtact gcgttgcta ctaacgtgc  
 240  
 tcaaaagtc ggtgacgtg taacttcta caaatctgt ttggtgcga tggagtcgg  
 300  
 tcaattctt taccctaagc gtaacttgac caagagcttc ctaagtctt cctctcgag  
 360  
 ctttaactg cggcgcttc cttcgttgt tggagagttt cccctccc tggttttct  
 420  
 actcggaat cggagagtc ggggtgact tttctctg gaactaagg tctggaagc  
 480  
 gccgttgcg aagctgttg ccgcggagt gtgaaagtgg aggttaagg ggcagaagt  
 540  
 gaactggat tcaaggaaa agttacggat ccttttggt tcaattgat cttcgaggg  
 600  
 aagaagacc tgaaccga cgaagacaa gaggtttaga actcgtctc ggaactaat  
 660  
 cttcgtcgt tctgaacaa aaaaataa tctattacc tggttatc cttttattt  
 720  
 gcttttgg gaactcggta acatgaac cggatcgagt agctaatga ctttcgatt  
 780  
 ctaattcac ccttttgc aaacttaat cttggtgaa ttgcattta arctcggct  
 840

(2) INFORMATION FOR SEQ ID NO:81:

630

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 148 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..148

(D) OTHER INFORMATION: / Ceres Seq. ID 1442748

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

His Glu Thr Ile Pro Leu Leu Arg Phe Leu His Lys Tyr Leu Thr His  
 1 10  
 Arg His Tyr Ile Val Leu Ser His Ala Ile Phe Ile Ser Leu Leu Thr  
 15 20  
 Val Asn Glu Ser Leu Phe Ser Lys Phe Tyr Phe Leu Ser Ala Met Ala  
 25 30  
 Gln Glu Asp Val Thr Ala Val Ala Thr Asn Gly Ala Gly Pro Val Glu  
 35 40  
 Thr His Leu Val Phe Thr Glu Phe Lys Gln Met Leu Leu Val Glu Ala  
 45 50  
 Gln Lys Val Gly Asp Ala Val Thr Phe Tyr Lys Ser Ala Phe Gly Ala  
 55 60  
 Ile Glu Ser Gly His Ser Leu Tyr Pro Lys Arg Asn Leu Thr Lys Ser  
 65 70  
 Phe Leu Met Phe Ser Leu Ser Leu Ile Ser Leu Ala Leu Pro Ser  
 75 80  
 Leu Phe Ala Thr Phe Pro Leu Ser Leu Val Phe Leu Leu Arg Asn Arg  
 85 90  
 Lys Val Arg Glu  
 145

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 102 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..102

(D) OTHER INFORMATION: / Ceres Seq. ID 1442749

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

Met Ala Gln Glu Asp Val Thr Ala Val Ala Thr Asn Gly Ala Gly Pro  
 1 10  
 Val Glu Thr His Leu Val Phe Thr Glu Phe Lys Gln Met Leu Leu Val  
 15 20  
 Glu Ala Gln Lys Val Gly Asp Ala Val Thr Phe Tyr Lys Ser Ala Phe  
 25 30  
 Gly Ala Ile Glu Ser Gly His Ser Leu Tyr Pro Lys Arg Asn Leu Thr  
 35 40  
 Lys Ser Phe Leu Met Phe Ser Leu Ser Leu Ile Ser Leu Ala Leu  
 45 50  
 Pro Ser Leu Phe Ala Thr Phe Pro Leu Ser Leu Val Phe Leu Leu Arg  
 55 60  
 Asn Arg Lys Val Arg Glu  
 100

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 74 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

631

(D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..74  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1442750  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:  
 Met Leu Leu Val Gly Ala Gln Lys Val Gly Asp Ala Val Thr Phe Tyr  
 1 5 10 15  
 Lys Ser Ala Phe Gly Ala Ile Glu Ser Gly His Ser Leu Tyr Pro Lys  
 20 25 30  
 Arg Asn Leu Thr Lys Ser Phe Leu Met Phe Ser Leu Leu Ser Leu Ile  
 35 40 45  
 Ser Leu Ala Leu Pro Ser Leu Phe Ala Thr Phe Pro Leu Ser Leu Val  
 50 55 60  
 Phe Leu Leu Arg Asn Arg Lys Val Arg Glu

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 513 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..513  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1459199

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

aactggatc ttaacacaga gaagagcac aatcggaaga aagagaaga caaagatgc 60  
 gtggatagg ctgaagcga gattcgctt cggatcatc ggggagatgc tctgatatc 120  
 gggcaattc cagctatca tccacaagc gttcatggc ggtcctaagc acatcgcca 180  
 cgacaaagg gacttgcta tggaaagac gacaagaaa gtcctcagc aactcgcca 240  
 tctcttcta gtatcgctt tatctctt gtgtctca ggtgcacat tctgtgaca 300  
 aataaagtc attcagaag aagaagaagc tgggggatc agtacttca tccacttg 360  
 attctctg gacatttaa agctttcaga aatcagacc caataacat tggttatca 420  
 attctctc attcgata ttgtgatc ttttaagtg catgatacac tgaactacc 480  
 tctgtatg tgaatgaaa aatgaatc tgc

(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 83 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..83  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1459200

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

Thr Trp Ile Leu Thr Gln Arg Arg Ser Thr Ile Gly Arg Lys Glu Lys  
 1 5 10 15  
 Thr Lys Met Ser Leu Val Trp Leu Glu Ala Met Leu Pro Leu Gly Ile  
 20 25 30  
 Ile Gly Gly Met Leu Cys Ile Met Gly Asn Ser Gln Tyr Tyr Ile His  
 35 40 45  
 Lys Ala Tyr His Gly Arg Pro Lys His Ile Gly His Asp Glu Trp Asp  
 50 55 60  
 Val Ala Met Glu Arg Arg Asp Lys Lys Val Val Glu Lys Ala Ala  
 65 70 75 80  
 Pro Ser Ser

632

(2) INFORMATION FOR SEQ ID NO:86:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 65 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..65  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1459201  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:  
 Met Ser Leu Val Trp Leu Glu Ala Met Leu Pro Leu Gly Ile Ile Gly  
 1 5 10 15  
 Gly Met Leu Cys Ile Met Gly Asn Ser Gln Tyr Tyr Ile His Lys Ala  
 20 25 30  
 Tyr His Gly Arg Pro Lys His Ile Gly His Asp Glu Trp Asp Val Ala  
 35 40 45  
 Met Glu Arg Arg Asp Lys Lys Val Val Glu Lys Ala Ala Pro Ser  
 50 55 60  
 Ser

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 57 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..57  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1459202

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

Met Leu Pro Leu Gly Ile Ile Gly Gly Met Leu Cys Ile Met Gly Asn  
 1 5 10 15  
 Ser Gln Tyr Tyr Ile His Lys Ala Tyr His Gly Arg Pro Lys His Ile  
 20 25 30  
 Gly His Asp Glu Trp Asp Val Ala Met Glu Arg Arg Asp Lys Lys Val  
 35 40 45  
 Val Glu Lys Ala Ala Pro Ser Ser

(2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1140 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..1140  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1565605

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

atgattgtt tacattttt cgttggttc ttgattttt tcttgatca aaactagtag 60  
 aaatttcca atttcttct cgttggttc tagtagtag tatttgaa ttgttagt 120  
 aggtagaaga agatgggac gacttagat gatacaag cagactagc atttgtagt 180  
 atgatttga acaaaagcaga ggaagagat agttagtga gactatata gtatggttc 240  
 agdttctga ggggggaca actcgactt actcaaatg ttgacaatc tactagttc 300  
 gcaagaagaag tctctgctt tttaagttt gtgaagact tgaatggtc tatcagctc 360  
 gtgcctaaag gaactctct tctcttgtt ttacttgga agtcaagaag cgcacttta 420  
 tctacatct tgttctctga tcaattgtc tggcttgga gatacgaat atataaagc 480

633

5 aaagacagag ctgagttact tggacgtata tctctctct gctggatggg atctcttgc  
 tgcacact tagcaggt tagtgatg ggaaggttt cttcatcaat gaagagatc  
 gaaagggac caagaatgg aaacaagtat cagatgagg attatcgtc taagtataa  
 aatcaaacg agaggtaact tgccttgatc aaacagctca tgcacatgt tttagcgtc  
 780 ggtctcttc agttagctcc acagagatc actctctgtg tccagggagc tttagcgtc  
 atcactcca tctctcttg ttaccagtgt ettcagcac gcccaagat caaacaccc  
 tgaagtaat cgaggaagct ggtgttaagg agaaagcat tccagaagt tictgtctt  
 900 cttattatca agaaatcca cgcgctat aaacatctg attaatct ttttgagg  
 aattcaggt attataat tagtccag attctcgata cacaagttc aaactcaac  
 1020 agtatgcat tattataga tctgttct cctacgaat aactagagc aaactcaac  
 1080 tacatatg tttagatg attattata tagactttg gctcatccc gctaaagct  
 1140

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 236 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ix) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..236

(2) OTHER INFORMATION: / Ceres Seq. ID 1565606

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

Met Gly Thr Thr Leu Asp Val Ser Arg Ala Glu Leu Ala Leu Val Val  
 1 5 10 15  
 Met Tyr Leu Asn Lys Ala Glu Ala Arg Asp Lys Leu Cys Arg Ala Ile  
 20 25 30 35  
 Gln Tyr Gly Ser Lys Phe Leu Ser Gly Gly Gln Pro Gly Thr Ala Gln  
 40 45  
 Asn Val Asp Lys Ser Thr Ser Leu Ala Arg Lys Val Phe Arg Leu Phe  
 50 55  
 Lys Phe Val Asn Asp Leu His Gly Ile Ser Pro Val Pro Lys Gly  
 60 65 70 75 80  
 Thr Pro Leu Pro Leu Val Leu Gly Lys Ser Lys Asn Ala Leu Leu  
 85 90 95  
 Ser Thr Phe Leu Phe Leu Asp Gln Ile Val Trp Leu Gly Arg Ser Gly  
 100 105 110  
 Ile Tyr Lys Asn Lys Glu Arg Ala Glu Leu Leu Gly Arg Ile Ser Leu  
 115 120 125  
 Phe Cys Trp Met Gly Ser Thr Val Cys Thr Thr Leu Val Glu Val Gly  
 130 135 140  
 Glu Met Gly Arg Leu Ser Ser Met Lys Lys Ile Glu Lys Gly Leu  
 145 150 155 160  
 Lys Asn Gly Asn Lys Tyr Gln Asp Glu Asp Tyr Arg Ala Lys Lys  
 165 170 175  
 Lys Ser Asn Glu Arg Ser Leu Ala Leu Ile Lys Ser Ala Met Asp Ile  
 180 185 190  
 Val Val Ala Ala Gly Leu Leu Gln Leu Ala Pro Thr Lys Ile Thr Pro  
 195 200 205  
 Arg Val Thr Gly Ala Phe Gly Phe Ile Thr Ser Ile Ile Ser Cys Tyr  
 210 215 220  
 Gln Leu Leu Pro Thr Arg Pro Lys Ile Lys Thr Pro  
 225 230 235

(2) INFORMATION FOR SEQ ID NO:90:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 220 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ix) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..132

634

(B) LOCATION: 1..220  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1565607

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

Met Tyr Leu Asn Lys Ala Glu Ala Arg Asp Lys Leu Cys Arg Ala Ile  
 1 5 10 15  
 Gln Tyr Gly Ser Lys Phe Leu Ser Gly Gly Gln Pro Gly Thr Ala 31n  
 20 25 30  
 Asn Val Asp Lys Ser Thr Ser Leu Ala Arg Lys Val Phe Arg Leu Phe  
 35 40 45  
 Lys Phe Val Asn Asp Leu His Gly Leu Ile Ser Pro Val Pro Lys Gly  
 50 55 60  
 Thr Pro Leu Pro Leu Val Leu Leu Gly Lys Ser Lys Asn Ala Leu Leu  
 65 70 75 80  
 Ser Thr Phe Leu Phe Leu Asp Gln Ile Val Trp Leu Gly Arg Ser Gly  
 85 90 95  
 Ile Tyr Lys Asn Lys Glu Arg Ala Glu Leu Leu Gly Arg Ile Ser Leu  
 100 105 110  
 Phe Cys Trp Met Gly Ser Thr Val Cys Thr Thr Leu Val Glu Val Gly  
 115 120 125  
 Glu Met Gly Arg Leu Ser Ser Ser Met Lys Lys Ile Glu Lys Gly Leu  
 130 135 140  
 Lys Asn Gly Asn Lys Tyr Gln Asp Glu Asp Tyr Arg Ala Lys Leu Lys  
 145 150 155 160  
 Lys Ser Asn Glu Arg Ser Leu Ala Leu Ile Lys Ser Ala Met Asp Ile  
 165 170 175  
 Val Val Ala Ala Gly Leu Leu Gln Leu Ala Pro Thr Lys Ile Thr Pro  
 180 185 190  
 Arg Val Thr Gly Ala Phe Gly Phe Ile Thr Ser Ile Ile Ser Cys Tyr  
 195 200 205  
 Gln Leu Leu Pro Thr Arg Pro Lys Ile Lys Thr Pro  
 210 215 220

(2) INFORMATION FOR SEQ ID NO:91:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 533 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ix) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..533  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1566686

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

atgattggg gttcttcgc cgaatttaa ttcacagat cgagagagaa aaactctct  
 60 cctctctc cgttcacgt cgcactcgg cctcacgta catttggtt agcaactctg  
 120 tatttagg attccaagt ctggcgaga ggaataaaa tcttccatg ctcagcttc  
 180 gttccaat tttctgac ttgttgatc ttgttgatc ttgttgatc ttgttgatc  
 240 tggggagca cgaatgagc ttgatctga ttggatata tagagaag agttgagatt  
 300 gtcagtggaa gcgctgcga agatcgctga tctaacac aacctgaa caaacacaa  
 360 atactgtt gatatttg ggcagctca cctctcga gctcagtg ttgtcaattg  
 420 catgacctt gggcgacaa ttgtgcag aggtttgtt cctcatctg aaatgcat  
 480 gggaaagga agaaagctc gtggcaag actaccaga gcaacacgc tgc

635  
(D) OTHER INFORMATION: / Ceres Seq. ID 1566687  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:  
Met Ser Gly Ala Glu Asp Asn Lys Ser Ser His Ala Gln Leu Ser Ser  
1 5 10 15  
Gln Ile Phe Leu Asp Leu Val Asp Ser Val Ile Ala Asp Val Ala Ser  
20 25 30  
Glu Cys His Arg Val Ala Arg Leu Gly Leu Asp Arg Asp Leu Asp Ile  
35 40 45  
Val Glu Glu Glu Leu Arg Leu Ser Val Glu Ala Arg Ala Lys Ile Ala  
50 55 60  
Asp Pro Ser Asn Asn Leu Glu Thr Asn Thr Lys Tyr Val Val Asp Ile  
65 70 75 80  
Phe Gly Gln Thr His Pro Val Ala Ser Glu Val Phe Asn Cys Met  
85 90 95  
15 Asn Cys Gly Arg Gln Ile Val Ala Gly Arg Phe Ala Pro His Leu Glu  
100 105 110  
Lys Cys Met Gly Lys Gly Arg Lys Ala Arg Gly Gln Gly Leu Pro Glu  
115 120 125  
Ala Gln Arg Leu  
130  
(2) INFORMATION FOR SEQ ID NO:93:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 815 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: ENA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..815  
(D) OTHER INFORMATION: / Ceres Seq. ID 1567367  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:  
gaacaaagac acgaagaagt cctcaaaaga atcattttca tgcctgcct tgcattttc  
1 tcaatggct gtcgtcgcgc cctcaaatgc gctccgcgcg gctcgcgcgc agacacattc  
60 tctctccat ccatcttcc cccgttttgc cgggttttcc tccaccgga aatcaccagc  
120 aattttccc gtcgtagcta tgcgtacctc gaataaagggtg aacaaatag atgcacaagt  
180 gaagaacaa tgcgtagcta tgcgtacctc ttttcaagggtg agtgacaaa taacagctga  
240 tgttttcaag aagctggaga agcaaaagt gttgagcaac gttgagaat ctggcctgct  
300 gtaaaagag aggggttggg accaaatggt ccatcttttg agaaattaa agtcttctcc  
360 aaagcagag accttggctc tctcagttcc cttagaact tagctggamg AACatgcctc  
420 gcggtcttag cctcgcgtgc attaccagct ctacagcgc ctatgtagc ctggtgttg  
480 atcccaatg actcaactac tctagtggtt gctcaggcgg ttctggccgg tgccttggg  
540 cttaacaggg ttgttttgtt ggttggttct gttgttttg atggacttca agaaactgac  
600 tgcattcttc tctgtaaac aaataaacc ccatgcttg tccaattgat tttgtcagt  
660 tgcgtattta tgcgtgatg gttcagttgt ctatggttag tccaagaat aagctgagt  
720 atagaagaag gctttataat aattaacaa atatt  
780  
(2) INFORMATION FOR SEQ ID NO:94:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 76 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..76  
(D) OTHER INFORMATION: / Ceres Seq. ID 1567368  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:  
Glu Thr Arg His Glu Ile Ser Lys Glu Ile Ile Phe Ile Ala Ser  
1 5 10 15  
Leu Arg Phe Phe Phe Asn Gly Cys Arg Arg Arg Ser Asn Ile Val Ser  
20 25 30

636  
Gly Gly Ser Ala Ala Asp Thr Ile Ser Leu Gln Ser His Ser Pro Pro  
35 40 45  
Leu Ser Pro Val Phe Leu His Arg Lys Ile Thr Ser Asn Phe Leu Arg  
50 55 60  
Arg Ser Tyr Gly Ser Pro Glu Lys Gly Glu Gln Ile  
65 70 75  
(2) INFORMATION FOR SEQ ID NO:95:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 120 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..120  
(D) OTHER INFORMATION: / Ceres Seq. ID 1567369  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:  
Lys Gln Asp Thr Lys Glu Ser Leu Lys Lys Ser Phe Ser Ser Ser Leu Arg  
1 5 10  
Phe Asp Phe Ser Ser Met Ala Val Val Gly Ala Pro Ile Ser Ser Pro  
15 20 25  
Ala Ala Gln Leu Gln Thr Gln Phe Leu Ser Asn Pro Ile Leu Pro Arg  
30 35 40 45  
Phe Arg Arg Ser Phe Ser Thr Gly Lys Ser Pro Ala Thr Phe Ser Val  
50 55 60  
Val Ala Met Ala Pro Gln Lys Lys Val Asn Lys Tyr Asp Ala Lys Trp  
65 70 75 80  
Lys Lys Gln Trp Tyr Ala Gly Leu Phe Phe Glu Gly Ser Glu Gln  
85 90 95  
Ile Asn Val Asp Val Phe Lys Lys Leu Glu Lys Arg Lys Val Leu Ser  
100 105 110  
Asn Val Glu Lys Ser Gly Leu Leu  
115 120  
(2) INFORMATION FOR SEQ ID NO:96:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 99 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..99  
(D) OTHER INFORMATION: / Ceres Seq. ID 1567370  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:  
Met Ala Val Val Gly Ala Pro Ile Ser Ser Pro Ala Ala Gln Leu Gln  
1 5 10  
Thr Gln Phe Leu Ser Asn Pro Ile Leu Pro Arg Phe Arg Arg Ser Phe  
15 20 25 30  
Ser Thr Gly Lys Ser Pro Ala Thr Phe Ser Val Val Ala Met Ala Pro  
35 40 45  
Gln Lys Lys Val Asn Lys Tyr Asp Ala Lys Trp Lys Lys Gln Trp Tyr  
50 55 60  
Gly Ala Gly Leu Phe Phe Glu Gly Ser Glu Gln Ile Asn Val Asp Val  
65 70 75 80  
Phe Lys Lys Leu Glu Lys Arg Lys Val Leu Ser Asn Val Glu Lys Ser  
85 90 95  
Gly Leu Leu  
100

637

- (A) LENGTH: 541 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:

5

(A) NAME/KEY: -

(B) LOCATION: 1..541

(D) OTHER INFORMATION: / Ceres Seq. ID 1570101

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

aaacattac gtgtgtatta tctgtttca cattttcta accctcttg gtaaatcgt 60  
 catgaagc caattatcg ccggtgcaat ctgatatct ccccgctgg aaaaaatggc 120  
 catgtgtca aagctctct tcaagtctac ttgtacttc gattctaac tctcttcaa 180  
 agtttcacg gtaccagag gagctccgtt caccgctgt ctgaattcg cagcagaga 240  
 gtcaaggtt cctccaaa cccagcct catcctat gatggatcg ggtcaatcc 300  
 tcaacaggt gaagaaacg ttttttcaa gcacgatct gaactaat tgcctctcg 360  
 tgaagagt gaagctgtg ttgtatgga tccatagat atgtatagc aactattaa 420  
 aactcagtc caaaaaaaa acataaata aactgaat gcttatgat caacttctt 480  
 gaataaat gaacgcga catctttta cggattgaag tgaattgat ttgtcaagt 540

(2) INFORMATION FOR SEQ ID NO:98:  
 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 93 amino acids  
 (B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..93

(D) OTHER INFORMATION: / Ceres Seq. ID 1570102

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

Met Ala Thr Gly Gly Lys Val Ser Phe Lys Val Thr Leu Thr Ser Asp 1  
 Pro Lys Leu Pro Phe Lys Val Phe Ser Val Pro Glu Gly Ala Phe 15  
 Thr Ala Val Leu Lys Phe Ala Ala Glu Glu Phe Lys Val Pro Pro Gln 20  
 Thr Ser Ala Ile Ile Thr Asn Asp Gly Ile Gly Ile Asn Pro Gln Gln 35  
 Ser Ala Gly Asn Val Phe Leu Lys His Gly Ser Glu Leu Arg Ile 40  
 Pro Arg Asp Arg Val Gly Ala Val Phe Val Met Asp Pro 80  
 90

(2) INFORMATION FOR SEQ ID NO:99:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 418 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..418

(D) OTHER INFORMATION: / Ceres Seq. ID 1571051

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

aaagaaga caacgatgca gaggtttaag tcaaacgaag ttgcgggtt ggcagtcggg 60  
 gctccttc ttggcgccac catcgctct cctaaagttg atgctttat cgcgtctct 120  
 cagagaagt cctggccat gtgcgaaa ttgtgagatc ttaagaagt agcttgaggc 180  
 cgttgaag gaacaggac aatcaatca ggaagattct ttgtttcag tgaactaca 240  
 aacacaagt cagtggttg gataaactc caagcaatg ttgtttccc ttgcctcga 300  
 tgcacaaat ctgaccatt tctcggtat ttatagttg ttctatttc ttgacacat 360

638

gataagtga atcggtccat tggtaatggt aatgttaaag ttgaagaatg tcttgttt  
 (2) INFORMATION FOR SEQ ID NO:100:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 104 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..104

(D) OTHER INFORMATION: / Ceres Seq. ID 1571052

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

Lys Arg Lys Thr Thr Met Gln Ser Leu Thr Ser Asn Glu Val Ala Gly 1  
 Leu Ala Val Gly Ala Leu Leu Gly Ala Thr Ile Ala Ala Pro Lys 15  
 Val Asp Ala Phe Ile Ala Ala Ser Gln Arg Arg Ser Leu Gly Met Cys 20  
 Arg Lys Cys Gly Asp Leu Lys Asn Val Ala Cys Gly Arg Cys Lys Gly 25  
 Thr Gly Thr Ile Lys Ser Gly Gly Phe Phe Gly Phe Ser Asp Ser 30  
 Asn Thr Arg Ser Val Ala Cys Asp Asn Cys Gln Ala Lys Gly Cys Phe 35  
 Pro Cys Pro Glu Cys Ser Lys Ser 40  
 100

(2) INFORMATION FOR SEQ ID NO:101:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 99 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..99

(D) OTHER INFORMATION: / Ceres Seq. ID 1571053

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

Met Gln Ser Leu Thr Ser Asn Glu Val Ala Gly Leu Ala Val Gly Ala 1  
 Leu Leu Leu Gly Ala Thr Ile Ala Ala Pro Lys Val Asp Ala Phe Ile 15  
 Ala Ala Ser Gln Arg Arg Ser Leu Gly Met Cys Arg Lys Cys Gly Asp 20  
 Leu Lys Asn Val Ala Cys Gly Arg Cys Lys Gly Thr Gly Thr Ile Lys 25  
 Ser Gly Gly Phe Phe Gly Phe Ser Asp Ser Ser Asn Thr Arg Ser Val 30  
 Ala Cys Asp Asn Cys Gln Ala Lys Gly Cys Phe Pro Cys Pro Glu Cys 35  
 Ser Lys Ser 40  
 90

(2) INFORMATION FOR SEQ ID NO:102:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 58 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide



639

(B) LOCATION: 1..58  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1571054  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:  
 Met Cys Arg Lys Cys Gly Asp Leu Lys Asn Val Ala Cys Gly Arg Cys  
 1 10  
 Lys Gly Thr Gly Thr Ile Lys Ser Gly Gly Phe Phe Gly Phe Ser Asp  
 20 30  
 Ser Ser Asn Thr Arg Ser Val Ala Cys Asp Asn Cys Gln Ala Lys Gly  
 35 45  
 Cys Phe Pro Cys Pro Glu Cys Ser Lys Ser  
 50 55

(2) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 444 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..444  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1571100

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

aactcaatt tgaattctt ggaagactt tgaattttc tctctcagt tacagctcgc 60  
 gatccagatc catatcagat catagcggt gtagtagcaa gaattgtga aaagtgtgt 120  
 aaactgata ttgtgtatc gatcgtgtg acatacatt cagctctag tctcctaata 180  
 gtaactaga atagtgagct tgcctgac agagaggaga agagctttt gaaggaggg 240  
 ttgtatcac tgaagctag agaatcaac agagatcagt atcccaaaag tgtgttagg 300  
 ctgtgtttt tgtgtttt tctacaatt taattttta ttctctttt ccagaaatta 360  
 accaactca tgaacagaga tgaattatc ttgctctct aagactcag atgattcgt 420  
 attgataaaa cttcatatatt tscoc

(2) INFORMATION FOR SEQ ID NO:104:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 93 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..93  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1571101

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

Asn Leu Leu Leu Ser Ser Leu Glu Gln Leu Trp Ser Phe Leu Ser Ser  
 1 10  
 Val Thr Ala Arg Asp Pro Asp Pro Tyr Pro Val Met Ala Gly Val Val  
 20 30  
 Ala Arg Ile Val Glu Lys Phe Gly Lys Arg Asp Ile Trp Tyr Arg Ile  
 35 40  
 Ser Leu Ile Ser Ile Pro Val Ser Ser Val Leu Ile Val Leu Gly Asn  
 50 60  
 Met Glu Leu Ala Arg Ile Arg Glu Glu Lys Lys Leu Lys Glu Arg  
 65 80  
 Phe Asp Gln Leu Arg Ala Arg Gly Ile Ile Arg Asp Gln  
 85 90

(2) INFORMATION FOR SEQ ID NO:105:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 66 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide

640

(ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..66  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1571102  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:  
 Met Ala Gly Val Val Ala Arg Ile Val Glu Lys Phe Gly Lys Arg Asp  
 1 10  
 Ile Trp Tyr Arg Ile Ser Leu Ile Ser Ile Pro Val Ser Ser Val Leu  
 20 30  
 Ile Val Leu Gly Asn Met Glu Leu Ala Arg Ile Arg Glu Glu Lys Lys  
 35 45  
 Leu Leu Lys Glu Arg Phe Asp Gln Leu Arg Ala Arg Gly Ile Ile Arg  
 50 60  
 Asp Gln  
 65

(2) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 636 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..636  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1665272

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

cagaataaa actattctt tgtttgtg ttatattac ttaaaaaa aaagagatc 60  
 agttggag aaagcagcaa adaaagaaa tggcagcaac atttgaaca ccatcagcg 120  
 tgaatgctt cagagatca tcatcacca ccaaacctt ctctcatcc ttttaaaac 180  
 caactaa ggcctgaagc acttgagac tgcgcgtgc atcggaggga agatcactt 240  
 gcttgagag caactgttg agagagatt tgaactgtt agatttggg ctgactgat 300  
 ggtagctcc gtgcagcatt ccagcgataa atgggaagag cctgacgggt ctctctcg 360  
 atagatcgg aactgagctc gctcacttcc cgcactctcc agctctcact tcaagttcg 420  
 ggttggtgt ggttcagcgg cacttaggct tctctctcg cctcacttc gacaaatcg 480  
 gattcaagg caggatcag gattactct aagataaact attctgttt tctgtttac 540  
 tatagctct ccttgggta tgtgaatat tatcaatca aacaaagctt tcttgcgat 600  
 ttgatgttc attctatc tcgacagtc ttctt

(2) INFORMATION FOR SEQ ID NO:107:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 140 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..140  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1665273

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

Met Ala Ala Thr Phe Ala Thr Pro Ser Thr Val Ile Gly Leu Gly Gly  
 1 10  
 Ser Ser Ile Thr Thr Lys Pro Phe Ser Ser Ser Phe Leu Lys Pro Thr  
 20 30  
 Leu Ser Ala Lys Asn Pro Leu Arg Leu Ala Gly Ala Ser Gly Gly Arg  
 35 45  
 Val Thr Cys Phe Glu Arg Asn Trp Leu Arg Arg Asp Leu Asn Val Val  
 50 60  
 Gly Phe Gly Leu Ile Gly Trp Leu Ala Pro Ser Ser Ile Pro Ala Ile  
 65 80  
 Asn Gly Lys Ser Leu Thr Gly Leu Phe Phe Asp Ser Ile Gly Thr Glu  
 85 95

641

Leu Ala His Phe Thr Pro Pro Ala Leu Thr Ser Gln Phe Thr Leu  
100 105 110  
Trp Leu Val Thr Trp His Leu Gly Leu Phe Leu Cys Leu Thr Phe Gly  
115 120  
Gln Ile Gly Phe Lys Gly Arg Thr Glu Asp Tyr Phe  
130 135 140

(2) INFORMATION FOR SEQ ID NO:108:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 522 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..522

(D) OTHER INFORMATION: / Ceres Seq. ID 1713895

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

aacttgat cttacacag agaagaaga caatcggaag aagaagaag acgaagatg 60  
cctggatg cctggaagc atgtgctc tcggaatcat cgtgggatg cttgtatca 120  
-ggcaattc tcaatactac atccacaag cttatcatg ccgtccaaag cacatcgcc 180  
agatgaatg ggaatgct atggaagac ggaagaaga agtcctcgaag aaagctgag 240  
ctctcttc atgattgct ttatctct ttgtcttc aggggctaa gtagacttca 300  
ctgtgaca aataagtc atccagaag aagaagaag cggggatct agtacttca 360  
tccaatg attctctg gacatattaa agcttcaga aatcagact caataactt 420  
tgattaca atattctt atcgggata tttgtatc ttttaattg catgataac 480  
tgaactac tcttgctg tgaatgaata aatgaatac tg

(2) INFORMATION FOR SEQ ID NO:109:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 83 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..83

(D) OTHER INFORMATION: / Ceres Seq. ID 1713896

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

Thr Trp Ile Leu Thr Gln Arg Arg Ser Thr Ile Gly Arg Lys Glu Lys  
1 10 15  
Thr Lys Met Ser Leu Val Trp Leu Glu Ala Met Leu Pro Leu Gly Ile  
20 25 30

Ile Gly Gly Met Leu Cys Ile Met Gly Asn Ser Gln Tyr Tyr Ile His

Lys Ala Tyr His Gly Arg Pro Lys His Ile Gly His Asp Glu Trp Asp

Val Ala Met Glu Arg Asp Lys Lys Val Val Glu Lys Ala Ala Ala

Pro Ser Ser

50

55

60

65

70

75

80

85

90

95

100

105

110

115

120

125

130

135

140

145

150

155

160

165

170

642

Met Ser Leu Val Trp Leu Glu Ala Met Leu Pro Leu Gly Ile Ile Gly  
1 5 10  
Gly Met Leu Cys Ile Met Gly Asn Ser Gln Tyr Tyr Ile His Lys Ala  
20 25 30  
Tyr His Gly Arg Pro Lys His Ile Gly His Asp Glu Trp Asp Val Ala  
35 40 45  
Met Glu Arg Arg Asp Lys Lys Val Val Glu Lys Ala Ala Pro Ser  
50 55 60

(2) INFORMATION FOR SEQ ID NO:111:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 57 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..57

(D) OTHER INFORMATION: / Ceres Seq. ID 1713898

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

Met Leu Pro Leu Gly Ile Ile Gly Met Leu Cys Ile Met Gly Asn  
1 5 10 15  
Ser Gln Tyr Tyr Ile His Lys Ala Tyr His Gly Arg Pro Lys His Ile  
20 25 30  
Gly His Asp Glu Trp Asp Val Ala Met Glu Arg Arg Asp Lys Lys Val  
35 40 45

(2) INFORMATION FOR SEQ ID NO:112:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 745 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..745

(D) OTHER INFORMATION: / Ceres Seq. ID 1923752

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

acgcttaga ttaaggact atcgttctg gttggatt ctcgccacc gcttgccta 60  
cgatcacac accacgtcc ttggatcag tctcaactt aagctctgat tcaatcac 120  
accaaagcty acttgcttg acttgcttc tcaaggatt gaggctgat gcaatcac 180  
aacaaatga agacaactt tcttattt attagaatc tactcaaac taatcttc 240  
taggtctta gaacaactt ctaagcttc tcaatctt tcttgagact tatggaaca 300  
ctctccatg tctctata tagaactgag ctaccctaa gctatatt catgaagct 360  
ctctgactc gcattctt ctatgaggt cctctgac tctaggaga tctcttga 420  
ctccatttt gagctttt ttgatcttc ttgagctt ttgagctt tagctctt gttattt 480  
cttctttc atagatgca tatataact aggaacaac acctcttc cctatgga 540  
ataggaatc ttacaactt tatcaactt gacattatt gatttttt ttcttggat 600  
tctctctt gaattctt tctgctatc ctttagcat ctttagcat ctttagcat 660  
gtctccaaag taagtgtat ttatcaaca cgcattgat ctttaagat cgtgggatt 720  
attatggc agagatgat actct

(2) INFORMATION FOR SEQ ID NO:113:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 35 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..65

(D) OTHER INFORMATION: / Ceres Seq. ID 1713897

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

643

(ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..35  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1923753  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:  
 Arg Leu Gly Phe Lys Asp Tyr Arg Ser Trp Leu Trp Ile Leu Ala His  
 1 5 10 15  
 Arg Leu Ala Tyr Glu Ser His Thr Val Leu Gly Ser Ser Leu Gln  
 20 25 30  
 Leu Lys Leu  
 35

(2) INFORMATION FOR SEQ ID NO:114:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 31 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..31  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1923754

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:114:  
 Met Glu Gln Leu Ser Ile Val Ser Leu Tyr Arg Thr Glu Leu Pro Leu  
 1 5 10 15  
 Ser Leu Tyr Leu Met Lys Ser Leu Ser Thr Arg Met Phe Ile Ser  
 20 25 30

(2) INFORMATION FOR SEQ ID NO:115:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 36 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..36  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1923755

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:  
 Met Arg Ser Leu Ser Thr Leu Arg Arg Ser Leu Leu Thr His Ile Leu  
 1 5 10 15  
 Ser Ser Leu Phe Asp Pro Leu Gly Cys Phe Ser Ser Leu Val Leu Ser  
 20 25 30  
 Ser Ser Ser  
 35

(2) INFORMATION FOR SEQ ID NO:116:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 487 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:  
 (A) NAME/KEY: -  
 (B) LOCATION: 1..487  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1976816

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:  
 aaacctagc aaaaactct cttctctt caattcatt cttctctac aatgcttgg  
 60  
 cgaacgag gatctgtgc tegtcttc gtctcgca ccgaagatc accactcc  
 120  
 cgttccca ccaaggcgt tctcgcttc cgtctctc aatctctt acctagcgt  
 180  
 cgttccact tctgtcacc ttccaggaat ctaggagccc taggttcac acagtcgtc  
 240

644

ttcctcgt acagtggtt ggtactct caactgacat ctaccttaa cgttaatttg  
 300  
 cgagttctt gtagctgc taagctact tgaagagca aaagagtgta attcaactgc  
 360  
 gcatgcgtga atcagctact tgaaggatg tgagatcgaa attgtgggat taccacaga  
 420  
 aaaaattgct taatgtgtg cctgttttag aagttttagt ttcaaatgt gggatttat  
 480  
 taattgt

(2) INFORMATION FOR SEQ ID NO:117:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 70 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..70  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1976817

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:  
 Asn Pro Ser Lys Asn Leu Ser Leu Ser Leu Asn Ser Leu Leu Leu Leu  
 1 5 10 15  
 Gln Trp Leu Gly Ala Thr Gln Asp Leu Leu Leu Val Leu Ser Ser Pro  
 20 25 30  
 Pro Pro Gln Asp His His Leu Ser Val Leu Pro Pro Arg Arg Phe Leu  
 35 40 45  
 Ala Ser Val Leu Leu Asn Pro Pro Tyr Leu Ala Val Ala Ser Pro Ser  
 50 55 60  
 Arg His Leu Pro Gly Ile  
 65 70

(2) INFORMATION FOR SEQ ID NO:118:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 57 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..57  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1976818

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:  
 Thr Leu Ala Lys Thr Ser Leu Ser Leu Ser Ile His Phe Ser Ser Tyr  
 1 5 10 15  
 Asn Gly Leu Ala Gln Arg Arg Ile Cys Cys Ser Phe Phe Arg Leu Arg  
 20 25 30  
 His Arg Lys Ile Thr Ile Ser Pro Phe Ser His His Gly Ala Ser Ser  
 35 40 45  
 Pro Pro Ser Ser Ile Leu Leu Thr  
 50 55

(2) INFORMATION FOR SEQ ID NO:119:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 93 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..93  
 (D) OTHER INFORMATION: / Ceres Seq. ID 1976819

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:  
 Met Ala Trp Arg Asn Ala Gly Ser Ala Ala Arg Ser Phe Val Ser Ala  
 1 5 10 15  
 Thr Ala Arg Ser Pro Ser Leu Arg Ser Thr Thr Ala Leu Pro Arg  
 20 25 30 35 40 45 50 55

645

20 25 30  
 35 40 45  
 50 55 60  
 65 70 75  
 80  
 85  
 90  
 95  
 100  
 105  
 110  
 115  
 120  
 125  
 130  
 135  
 140  
 145  
 150  
 155  
 160  
 165  
 170  
 175  
 180  
 185  
 190  
 195  
 200  
 205  
 210  
 215  
 220  
 225  
 230  
 235  
 240  
 245  
 250  
 255  
 260  
 265  
 270  
 275  
 280  
 285  
 290  
 295  
 300  
 305  
 310  
 315  
 320  
 325  
 330  
 335  
 340  
 345  
 350  
 355  
 360  
 365  
 370  
 375  
 380  
 385  
 390  
 395  
 400  
 405  
 410  
 415  
 420  
 425  
 430  
 435  
 440  
 445  
 450  
 455  
 460  
 465  
 470  
 475  
 480  
 485  
 490  
 495  
 500  
 505  
 510  
 515  
 520  
 525  
 530  
 535  
 540  
 545  
 550  
 555  
 560  
 565  
 570  
 575  
 580  
 585  
 590  
 595  
 600  
 605  
 610  
 615  
 620  
 625  
 630  
 635  
 640  
 645  
 650  
 655  
 660  
 665  
 670  
 675  
 680  
 685  
 690  
 695  
 700  
 705  
 710  
 715  
 720  
 725  
 730  
 735  
 740  
 745  
 750  
 755  
 760  
 765  
 770  
 775  
 780  
 785  
 790  
 795  
 800  
 805  
 810  
 815  
 820  
 825  
 830  
 835  
 840  
 845  
 850  
 855  
 860  
 865  
 870  
 875  
 880  
 885  
 890  
 895  
 900  
 905  
 910  
 915  
 920  
 925  
 930  
 935  
 940  
 945  
 950  
 955  
 960  
 965  
 970  
 975  
 980  
 985  
 990  
 995  
 1000  
 1005  
 1010  
 1015  
 1020  
 1025  
 1030  
 1035  
 1040  
 1045  
 1050  
 1055  
 1060  
 1065  
 1070  
 1075  
 1080  
 1085  
 1090  
 1095  
 1100  
 1105  
 1110  
 1115  
 1120  
 1125  
 1130  
 1135  
 1140  
 1145  
 1150  
 1155  
 1160  
 1165  
 1170  
 1175  
 1180  
 1185  
 1190  
 1195  
 1200  
 1205  
 1210  
 1215  
 1220  
 1225  
 1230  
 1235  
 1240  
 1245  
 1250  
 1255  
 1260  
 1265  
 1270  
 1275  
 1280  
 1285  
 1290  
 1295  
 1300  
 1305  
 1310  
 1315  
 1320  
 1325  
 1330  
 1335  
 1340  
 1345  
 1350  
 1355  
 1360  
 1365  
 1370  
 1375  
 1380  
 1385  
 1390  
 1395  
 1400  
 1405  
 1410  
 1415  
 1420  
 1425  
 1430  
 1435  
 1440  
 1445  
 1450  
 1455  
 1460  
 1465  
 1470  
 1475  
 1480  
 1485  
 1490  
 1495  
 1500  
 1505  
 1510  
 1515  
 1520  
 1525  
 1530  
 1535  
 1540  
 1545  
 1550  
 1555  
 1560  
 1565  
 1570  
 1575  
 1580  
 1585  
 1590  
 1595  
 1600  
 1605  
 1610  
 1615  
 1620  
 1625  
 1630  
 1635  
 1640  
 1645  
 1650  
 1655  
 1660  
 1665  
 1670  
 1675  
 1680  
 1685  
 1690  
 1695  
 1700  
 1705  
 1710  
 1715  
 1720  
 1725  
 1730  
 1735  
 1740  
 1745  
 1750  
 1755  
 1760  
 1765  
 1770  
 1775  
 1780  
 1785  
 1790  
 1795  
 1800  
 1805  
 1810  
 1815  
 1820  
 1825  
 1830  
 1835  
 1840  
 1845  
 1850  
 1855  
 1860  
 1865  
 1870  
 1875  
 1880  
 1885  
 1890  
 1895  
 1900  
 1905  
 1910  
 1915  
 1920  
 1925  
 1930  
 1935  
 1940  
 1945  
 1950  
 1955  
 1960  
 1965  
 1970  
 1975  
 1980  
 1985  
 1990  
 1995  
 2000  
 2005  
 2010  
 2015  
 2020  
 2025  
 2030  
 2035  
 2040  
 2045  
 2050  
 2055  
 2060  
 2065  
 2070  
 2075  
 2080  
 2085  
 2090  
 2095  
 2100  
 2105  
 2110  
 2115  
 2120  
 2125  
 2130  
 2135  
 2140  
 2145  
 2150  
 2155  
 2160  
 2165  
 2170  
 2175  
 2180  
 2185  
 2190  
 2195  
 2200  
 2205  
 2210  
 2215  
 2220  
 2225  
 2230  
 2235  
 2240  
 2245  
 2250  
 2255  
 2260  
 2265  
 2270  
 2275  
 2280  
 2285  
 2290  
 2295  
 2300  
 2305  
 2310  
 2315  
 2320  
 2325  
 2330  
 2335  
 2340  
 2345  
 2350  
 2355  
 2360  
 2365  
 2370  
 2375  
 2380  
 2385  
 2390  
 2395  
 2400  
 2405  
 2410  
 2415  
 2420  
 2425  
 2430  
 2435  
 2440  
 2445  
 2450  
 2455  
 2460  
 2465  
 2470  
 2475  
 2480  
 2485  
 2490  
 2495  
 2500  
 2505  
 2510  
 2515  
 2520  
 2525  
 2530  
 2535  
 2540  
 2545  
 2550  
 2555  
 2560  
 2565  
 2570  
 2575  
 2580  
 2585  
 2590  
 2595  
 2600  
 2605  
 2610  
 2615  
 2620  
 2625  
 2630  
 2635  
 2640  
 2645  
 2650  
 2655  
 2660  
 2665  
 2670  
 2675  
 2680  
 2685  
 2690  
 2695  
 2700  
 2705  
 2710  
 2715  
 2720  
 2725  
 2730  
 2735  
 2740  
 2745  
 2750  
 2755  
 2760  
 2765  
 2770  
 2775  
 2780  
 2785  
 2790  
 2795  
 2800  
 2805  
 2810  
 2815  
 2820  
 2825  
 2830  
 2835  
 2840  
 2845  
 2850  
 2855  
 2860  
 2865  
 2870  
 2875  
 2880  
 2885  
 2890  
 2895  
 2900  
 2905  
 2910  
 2915  
 2920  
 2925  
 2930  
 2935  
 2940  
 2945  
 2950  
 2955  
 2960  
 2965  
 2970  
 2975  
 2980  
 2985  
 2990  
 2995  
 3000  
 3005  
 3010  
 3015  
 3020  
 3025  
 3030  
 3035  
 3040  
 3045  
 3050  
 3055  
 3060  
 3065  
 3070  
 3075  
 3080  
 3085  
 3090  
 3095  
 3100  
 3105  
 3110  
 3115  
 3120  
 3125  
 3130  
 3135  
 3140  
 3145  
 3150  
 3155  
 3160  
 3165  
 3170  
 3175  
 3180  
 3185  
 3190  
 3195  
 3200  
 3205  
 3210  
 3215  
 3220  
 3225  
 3230  
 3235  
 3240  
 3245  
 3250  
 3255  
 3260  
 3265  
 3270  
 3275  
 3280  
 3285  
 3290  
 3295  
 3300  
 3305  
 3310  
 3315  
 3320  
 3325  
 3330  
 3335  
 3340  
 3345  
 3350  
 3355  
 3360  
 3365  
 3370  
 3375  
 3380  
 3385  
 3390  
 3395  
 3400  
 3405  
 3410  
 3415  
 3420  
 3425  
 3430  
 3435  
 3440  
 3445  
 3450  
 3455  
 3460  
 3465  
 3470  
 3475  
 3480  
 3485  
 3490  
 3495  
 3500  
 3505  
 3510  
 3515  
 3520  
 3525  
 3530  
 3535  
 3540  
 3545  
 3550  
 3555  
 3560  
 3565  
 3570  
 3575  
 3580  
 3585  
 3590  
 3595  
 3600  
 3605  
 3610  
 3615  
 3620  
 3625  
 3630  
 3635  
 3640  
 3645  
 3650  
 3655  
 3660  
 3665  
 3670  
 3675  
 3680  
 3685  
 3690  
 3695  
 3700  
 3705  
 3710  
 3715  
 3720  
 3725  
 3730  
 3735  
 3740  
 3745  
 3750  
 3755  
 3760  
 3765  
 3770  
 3775  
 3780  
 3785  
 3790  
 3795  
 3800  
 3805  
 3810  
 3815  
 3820  
 3825  
 3830  
 3835  
 3840  
 3845  
 3850  
 3855  
 3860  
 3865  
 3870  
 3875  
 3880  
 3885  
 3890  
 3895  
 3900  
 3905  
 3910  
 3915  
 3920  
 3925  
 3930  
 3935  
 3940  
 3945  
 3950  
 3955  
 3960  
 3965  
 3970  
 3975  
 3980  
 3985  
 3990  
 3995  
 4000  
 4005  
 4010  
 4015  
 4020  
 4025  
 4030  
 4035  
 4040  
 4045  
 4050  
 4055  
 4060  
 4065  
 4070  
 4075  
 4080  
 4085  
 4090  
 4095  
 4100  
 4105  
 4110  
 4115  
 4120  
 4125  
 4130  
 4135  
 4140  
 4145  
 4150  
 4155  
 4160  
 4165  
 4170  
 4175  
 4180  
 4185  
 4190  
 4195  
 4200  
 4205  
 4210  
 4215  
 4220  
 4225  
 4230  
 4235  
 4240  
 4245  
 4250  
 4255  
 4260  
 4265  
 4270  
 4275  
 4280  
 4285  
 4290  
 4295  
 4300  
 4305  
 4310  
 4315  
 4320  
 4325  
 4330  
 4335  
 4340  
 4345  
 4350  
 4355  
 4360  
 4365  
 4370  
 4375  
 4380  
 4385  
 4390  
 4395  
 4400  
 4405  
 4410  
 4415  
 4420  
 4425  
 4430  
 4435  
 4440  
 4445  
 4450  
 4455  
 4460  
 4465  
 4470  
 4475  
 4480  
 4485  
 4490  
 4495  
 4500  
 4505  
 4510  
 4515  
 4520  
 4525  
 4530  
 4535  
 4540  
 4545  
 4550  
 4555  
 4560  
 4565  
 4570  
 4575  
 4580  
 4585  
 4590  
 4595  
 4600  
 4605  
 4610  
 4615  
 4620  
 4625  
 4630  
 4635  
 4640  
 4645  
 4650  
 4655  
 4660  
 4665  
 4670  
 4675  
 4680  
 4685  
 4690  
 4695  
 4700  
 4705  
 4710  
 4715  
 4720  
 4725  
 4730  
 4735  
 4740  
 4745  
 4750  
 4755  
 4760  
 4765  
 4770  
 4775  
 4780  
 4785  
 4790  
 4795  
 4800  
 4805  
 4810  
 4815  
 4820  
 4825  
 4830  
 4835  
 4840  
 4845  
 4850  
 4855  
 4860  
 4865  
 4870  
 4875  
 4880  
 4885  
 4890  
 4895  
 4900  
 4905  
 4910  
 4915  
 4920  
 4925  
 4930  
 4935  
 4940  
 4945  
 4950  
 4955  
 4960  
 4965  
 4970  
 4975  
 4980  
 4985  
 4990  
 4995  
 5000  
 5005  
 5010  
 5015  
 5020  
 5025  
 5030  
 5035  
 5040  
 5045  
 5050  
 5055  
 5060  
 5065  
 5070  
 5075  
 5080  
 5085  
 5090  
 5095  
 5100  
 5105  
 5110  
 5115  
 5120  
 5125  
 5130  
 5135  
 5140  
 5145  
 5150  
 5155  
 5160  
 5165  
 5170  
 5175  
 5180  
 5185  
 5190  
 5195  
 5200  
 5205  
 5210  
 5215  
 5220  
 5225  
 5230  
 5235  
 5240  
 5245  
 5250  
 5255  
 5260  
 5265  
 5270  
 5275  
 5280  
 5285  
 5290  
 5295  
 5300  
 5305  
 5310  
 5315  
 5320  
 5325  
 5330  
 5335  
 5340  
 5345  
 5350  
 5355  
 5360  
 5365  
 5370  
 5375  
 5380  
 5385  
 5390  
 5395  
 5400  
 5405  
 5410  
 5415  
 5420  
 5425  
 5430  
 5435  
 5440  
 5445  
 5450  
 5455  
 5460  
 5465  
 5470  
 5475  
 5480  
 5485  
 5490  
 5495  
 5500  
 5505  
 5510  
 5515  
 5520  
 5525  
 5530  
 5535  
 5540  
 5545  
 5550  
 5555  
 5560  
 5565  
 5570  
 5575  
 5580  
 5585  
 5590  
 5595  
 5600  
 5605  
 5610  
 5615  
 5620  
 5625  
 5630  
 5635  
 5640  
 5645  
 5650  
 5655  
 5660  
 5665  
 5670  
 5675  
 5680  
 5685  
 5690  
 5695  
 5700  
 5705  
 5710  
 5715  
 5720  
 5725  
 5730  
 5735  
 5740  
 5745  
 5750  
 5755  
 5760  
 5765  
 5770  
 5775  
 5780  
 5785  
 5790  
 5795  
 5800  
 5805  
 5810  
 5815  
 5820  
 5825  
 5830  
 5835  
 5840  
 5845  
 5850  
 5855  
 5860  
 5865  
 5870  
 5875  
 5880  
 5885  
 5890  
 5895  
 5900  
 5905  
 5910  
 5915  
 5920  
 5925  
 5930  
 5935  
 5940  
 5945  
 5950  
 5955  
 5960  
 5965  
 5970  
 5975  
 5980  
 5985  
 5990  
 5995  
 6000  
 6005  
 6010  
 6015  
 6020  
 6025  
 6030  
 6035  
 6040  
 6045  
 6050  
 6055  
 6060  
 6065  
 6070  
 6075  
 6080  
 6085  
 6090  
 6095  
 6100  
 6105  
 6110  
 6115  
 6120  
 6125  
 6130  
 6135  
 6140  
 6145  
 6150  
 6155  
 6160  
 6165  
 6170  
 6175  
 6180  
 6185  
 6190  
 6195  
 6200  
 6205  
 6210  
 6215  
 6220  
 6225  
 6230  
 6235  
 6240  
 6245  
 6250  
 6255  
 6260  
 6265  
 6270  
 6275  
 6280  
 6285  
 6290  
 6295  
 6300  
 6305  
 6310  
 6315  
 6320  
 6325  
 6330  
 6335  
 6340  
 6345  
 6350  
 6355  
 6360  
 6365  
 6370  
 6375  
 6380  
 6385  
 6390  
 6395  
 6400  
 6405  
 6410  
 6415  
 6420  
 6425  
 6430  
 6435  
 6440  
 6445  
 6450  
 6455  
 6460  
 6465  
 6470  
 6475  
 6480  
 6485  
 6490  
 6495  
 6500  
 6505  
 6510  
 6515  
 6520  
 6525  
 6530  
 6535  
 6540  
 6545  
 6550  
 6555  
 6560  
 6565  
 6570  
 6575  
 6580  
 6585  
 6590  
 6595  
 6600  
 6605  
 6610  
 6615  
 6620  
 6625  
 6630  
 6635  
 6640  
 6645  
 6650  
 6655  
 6660  
 6665  
 6670  
 6675  
 6680  
 6685  
 6690  
 6695  
 6700  
 6705  
 6710  
 6715  
 6720  
 6725  
 6730  
 6735  
 6740  
 6745  
 6750  
 6755  
 6760  
 6765  
 6770  
 6775  
 6780  
 6785  
 6790  
 6795  
 6800  
 6805  
 6810  
 6815  
 6820  
 6825  
 6830  
 6835  
 6840  
 6845  
 6850  
 6855  
 6860  
 6865  
 6870  
 6875  
 6880  
 6885  
 6890  
 6895  
 6900  
 6905

- 647
- 20 Val Ile Ser His Ile Arg Leu Ala Ile Thr Ile Asn Ser Lys Leu 30  
35 45
- 25
- (2) INFORMATION FOR SEQ ID NO:126:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 33 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..33  
(D) OTHER INFORMATION: / Ceres Seq. ID 2025404  
Phe Phe Phe Phe Cys Glu Ile Glu Met Ile Ser His Lys Asn Ser Ile 15  
1 10  
Asp Asn Phe Thr Lys Glu Lys Lys Asn Leu Ser Ile Lys Glu Glu Glu Lys 30  
20 Asp
- (2) INFORMATION FOR SEQ ID NO:127:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1954 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..1954  
(D) OTHER INFORMATION: / Ceres Seq. ID 2025479  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:  
atgggaacctt atatacaacg aagaatata gttatcaag tttatcatgc agaagaaga 60  
gaagataca atgattgtga ataggacaa cttctcttg tcatattga tattatgct 120  
gtttcttcc agattgtga gcttgggt gaagccaaa agcgtaaag cgtgaacag 180  
cagcgttaa tagaattga atttttcaga actgtgtgt cgtatgtgt tctctacta 240  
aagccatggt tctctgatt gcaagaagcg atgtctgatt tccagtgtgt tccagagat 300  
caggaagaac agtcttgat gacgaacaac gaagaagaag atctgttga tgttgaggt 360  
cttgagccta tctaatgta gctcttagc tctctcttc cttcttgca ttgctggt 420  
gatcataag cagacaaggg tadacaactc tttctttga caccatacc cttaggaaaa 480  
ttgtaattcc taataaccca gaatgatctt aagcttacag ctaaaagat tttccagac 540  
actgtgcaa atgaaccgt tgaagtctt aaagaacaa ggtatgatgt tttagaggt 600  
gagtcgtga aaactgagg acttgtaa tcttagtcc atgcatgga tttcttgaa 660  
aatttggtt agtaaacgc atgttcttca ccagtttca ggagaagat tctgttgag 720  
cttttaaga caccttgctt gaagctgctg cctccgaat cgtatcaat gtttaacg 780  
gttctgaat ctctcaact ggaagaacaa ggggtttga agttactcg ctctgaatg 840  
ggatcatctg gattgaaa aacagaacat ctctcttga agttactga gctcttgga 900  
atacaaatg ctctgatac tagaataacg gatcttgat cctgcctgt tigtgttt 960  
tctctcta aacatgct tttcaatgag cctgtgatg agaagaacc aattgatga 1020  
actgtgatg cttttgat tccaactt ataccaga ataccaga caaacacac tacagagga 1080  
agcatgcaa ggtggttga gattcttca ttgtttaa agcagagag cataatgag 1140  
agaacagaa caaagcagg tgaagtact ctgaagaag agtatgag aagatttga 1200  
gaagcaaca tcatgacg tccgttcaac tcatgaca caacaacac agtgaagga 1260  
aataaaga atgtttctt ggaatgta gaagaatga gggccaatga ggaatgat 1320  
gggttaagt gcaggtttt ggtctcaaa ctctatcaag tactcatac ggaactgac 1380  
atacaaatg ctttttatg tctgtact ctattagct gctacatac tcaaatgac 1440  
aaacatac atatacala tagaagatc gaaaacaa tatctctc ctcaactca 1500  
caagaactc aactaacctt gctttctc ttgtaccaa gttctctc ctcaactca 1560  
ttaccccgag ttggagac aactctgc caagagac agtatgag cctgtgct 1620  
cgatgtct caactcact aacttgagc ttggggcacc agtatgag cctgtgct 1680  
ctattctct tggctcaat gatctcgat ttgcggttg ccttgacc gcctcaagc 1740
- 648
- 25 teagcttct tggcatcac atgacacac ctattcact taacttggct cttacagctt 1800  
gtgaggtac cttctctgt ggtatcctt gccacata gttacatat attcaactt 1860  
tctattatg ttttgattt ggtttgtct aacatcaaa ttcaacata aattcatga 1920  
taagcaata aagaataaa ttaagattt tatg
- (2) INFORMATION FOR SEQ ID NO:128:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 496 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..496  
(D) OTHER INFORMATION: / Ceres Seq. ID 2025480  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:  
Met Glu Pro Tyr Ile Thr Arg Arg Asp Thr Val Asn Gln Val Leu His 15  
1 5 10  
Ala Glu Glu Glu Glu Tyr Asn Asp Cys Glu Leu Asp Gln Leu Phe 20  
20 25  
Leu Val His Ser Asp Ile Arg Ser Val Leu Leu Gln Ile Asp Glu Leu 35  
35 40 45  
Val Val Glu Ala Thr Lys Arg Lys Thr Val Ser Lys His Gly Leu Ile 50  
50 55 60  
Glu Val Glu Ser Phe Arg Thr Val Leu Ser Asp Met Leu Ser Ser Leu 65  
65 70 75 80  
Lys Pro Trp Phe Pro Arg Leu Gln Ala Met Ser Asp Phe Gln Leu 85  
85 90 95  
Leu Pro Glu Asp Glu Glu Gln Ser Leu Met Ser Thr Asn Glu Glu 100  
100 105 110  
Glu Asp Leu Phe Asp Val Glu Ser Pro Glu Pro Thr Gln Phe Glu Pro 115  
115 120 125  
Leu Val Ser Pro Ser Pro Leu Val His Trp Arg Gly Asp His Asn Ala 130  
130 135 140  
Asp Lys Gly Arg Gln Leu Phe Leu Leu Thr Pro Leu Pro Leu Gly Lys 145  
145 150 155 160  
Ser Glu Phe Leu Lys His Gln Asn Ala Ser Lys Leu Thr Ala Lys Arg 165  
165 170 175  
Ile Phe Pro Asp Thr Val Ala Asn Glu Pro Leu Glu Ala Ser Lys Glu 180  
180 185 190  
Thr Ser Asp Asp Val Leu Gly Gly Glu Ser Leu Lys Thr Ala Gly Leu 195  
195 200 205  
Gly Lys Ser Leu Val His Ala Met Asp Phe Ser Glu Asn Leu Val Glu 210  
210 215 220  
Tyr Lys Pro Cys Ser Ser Pro Val Leu Arg Arg Lys Ile Leu Ser Glu 225  
225 230 235 240  
Leu Leu Met Thr Pro Cys Leu Lys Leu Ser Pro Pro Lys Ser Cys Thr 245  
245 250 255  
Met Phe Lys Pro Val Pro Glu Ser Ser Gln Leu Gly Lys Gln Gly Ala 260  
260 265 270  
Cys Lys Ser Thr Cys Ser Glu Leu Gly Ser Ser Gly Ile Glu Lys Thr 275  
275 280 285  
Asp Asn Leu Cys Ser Lys Tyr Pro Glu Leu Leu Gly Ile Gln His Ala 290  
290 295 300  
Pro Ile Thr Arg Lys Thr Asp Leu Glu Ser Ser Pro Val Trp Trp Phe 305  
305 310 315 320  
Ser Pro Pro Lys Thr Cys Val Leu Met Glu Pro Val Asn Glu Lys Lys 325  
325 330 335  
Pro Ile Asp Glu Thr Gly Gly Ser Phe Asp Val Pro Asn Ile Ile Pro 340  
340 345 350  
Glu Ala Lys His Thr Thr Glu Gly Ser Met Ser Met Val Val Glu Ser 355  
355 360 365

649

Thr Pro Leu Phe Lys Glu Pro Glu Ser Ile Met Thr Arg Asn Arg Thr  
370 375 380  
Lys Ala Gly Glu Ser Thr Leu Lys Lys Glu Leu Trp Thr Arg Phe Glu  
385 390 395 400  
Glu Ala Thr Ile His Asp Ser Arg Phe Asn Ser Met Thr Thr Thr  
405 410 415  
Thr Val Arg Gly Asn Asn Lys Lys Cys Phe Met Glu Met Leu Glu Glu  
420 425 430  
Val Ser Gly Asn Glu Glu Asp His Glu Leu Ser Gly Glu Phe Val Ala  
435 440 445  
Leu Lys Pro Tyr Cln Val Leu Ile Ser Glu Leu Ser Ile Thr Asn Ala  
450 455 460  
Phe His Cys Leu Ser Thr Leu Ile Ser Cys Tyr Ile Thr Gln Ser Asn  
465 470 475  
Lys His Thr His Ile Thr Tyr Arg Lys Ile Glu Asn Lys Arg Ser Gly  
480 485 490 495

{2} INFORMATION FOR SEQ ID NO:129:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 421 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(E) MOLECULE TYPE: peptide

(F) FEATURE:

(G) NAME/KEY: peptide

(H) LOCATION: 1...421

(I) OTHER INFORMATION: / Ceres Seq. ID 2025481

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

Met Leu Ser Ser Leu Lys Pro Trp Phe Pro Arg Leu Gln Glu Ala Met  
1 5 10 15  
Ser Asp Phe Gln Leu Leu Pro Glu Asp Gln Glu Gln Ser Leu Met  
20 25 30  
Ser Thr Asn Glu Glu Asp Leu Phe Asp Val Glu Ser Pro Glu Pro  
35 40 45

Thr Gln Phe Glu Pro Leu Val Ser Pro Ser Pro Leu Val His Trp Arg  
50 55 60  
Gly Asp His Asn Ala Asp Lys Gly Arg Gln Leu Phe Leu Thr Pro  
65 70 75 80  
Leu Pro Leu Gly Lys Ser Glu Phe Leu Lys His Gln Asn Ala Ser Lys  
85 90 95

Leu Thr Ala Lys Arg Ile Phe Pro Asp Thr Val Ala Asn Glu Pro Leu  
100 105 110  
Glu Ala Ser Lys Glu Thr Ser Asp Asp Val Leu Gly Gly Glu Ser Leu  
115 120 125  
Lys Thr Ala Gly Leu Gly Lys Ser Leu Val His Ala Met Asp Phe Ser  
130 135 140

Glu Asn Leu Val Glu Tyr Lys Pro Cys Ser Ser Pro Val Leu Arg Arg  
145 150 155 160  
Lys Ile Leu Ser Glu Leu Leu Met Thr Pro Cys Leu Lys Leu Ser Pro  
165 170 175  
Pro Lys Ser Cys Thr Met Phe Lys Pro Val Pro Glu Ser Ser Gln Leu  
180 185 190

Gly Lys Gln Gly Ala Cys Lys Ser Thr Cys Ser Glu Leu Gly Ser Ser  
195 200 205  
Gly Ile Glu Lys Thr Asp Asn Leu Cys Ser Lys Tyr Pro Glu Leu Leu  
210 215 220

Gly Ile Gln His Ala Pro Ile Thr Arg Lys Thr Asp Leu Glu Ser Ser  
225 230 235 240  
Pro Val Trp Trp Phe Ser Pro Pro Lys Thr Cys Val Leu Met Glu Pro  
245 250 255

Val Asn Glu Lys Lys Pro Ile Asp Glu Thr Gly Gly Ser Phe Asp Val  
260 265 270

650

Pro Asn Ile Ile Pro Glu Ala Lys His Thr Thr Glu Gly Ser Met Ser  
275 280 285  
Met Val Val Glu Ser Thr Pro Leu Phe Lys Glu Pro Glu Ser Ile Met  
290 295 300  
Thr Arg Asn Arg Thr Lys Ala Gly Glu Ser Thr Leu Lys Lys Glu Leu  
305 310 315 320  
Trp Thr Arg Phe Glu Glu Ala Thr Ile His Asp Ser Arg Phe Asn Ser  
325 330 335  
Met Thr Thr Thr Thr Val Arg Gly Asn Asn Lys Lys Cys Phe Met  
340 345 350  
Glu Met Leu Glu Glu Val Ser Gly Asn Glu Glu Asp His Glu Leu Ser  
355 360 365  
Gly Glu Phe Val Ala Leu Lys Pro Tyr Gln Val Leu Ile Ser Glu Leu  
370 375 380  
Ser Ile Thr Asn Ala Phe His Cys Leu Ser Thr Leu Ile Ser Cys Tyr  
385 390 395 400  
Ile Thr Gln Ser Asn Lys His Thr His Ile Thr Tyr Arg Lys Ile Glu  
405 410 415  
Asn Lys Arg Ser Gly

{2} INFORMATION FOR SEQ ID NO:130:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 406 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(E) MOLECULE TYPE: peptide

(F) FEATURE:

(G) NAME/KEY: peptide

(H) LOCATION: 1...406

(I) OTHER INFORMATION: / Ceres Seq. ID 2025482

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

Met Ser Asp Phe Gln Leu Leu Pro Glu Asp Gln Glu Gln Ser Leu  
1 5 10 15  
Met Ser Thr Asn Glu Glu Asp Leu Phe Asp Val Glu Ser Pro Glu  
20 25 30  
Pro Thr Gln Phe Glu Pro Leu Val Ser Pro Ser Pro Leu Val His Trp  
35 40 45  
Arg Gly Asp His Asn Ala Asp Lys Gly Arg Gln Leu Phe Leu Thr  
50 55 60  
Pro Leu Pro Leu Gly Lys Ser Glu Phe Leu Lys His Gln Asn Ala Ser  
65 70 75 80  
Lys Leu Thr Ala Lys Arg Ile Phe Pro Asp Thr Val Ala Asn Glu Pro  
85 90 95

Leu Glu Ala Ser Lys Glu Thr Ser Asp Asp Val Leu Gly Gly Glu Ser  
100 105 110  
Leu Lys Thr Ala Gly Leu Gly Lys Ser Leu Val His Ala Met Asp Phe  
115 120 125  
Ser Glu Asn Leu Val Glu Tyr Lys Pro Cys Ser Ser Pro Val Leu Arg  
130 135 140  
Arg Lys Ile Leu Ser Glu Leu Met Thr Pro Cys Leu Lys Leu Ser  
145 150 155 160  
Pro Pro Lys Ser Cys Thr Met Phe Lys Pro Val Pro Glu Ser Ser Gln  
165 170 175  
Leu Gly Lys Gln Gly Ala Cys Lys Ser Thr Cys Ser Glu Leu Gly Ser  
180 185 190

Ser Gly Ile Glu Lys Thr Asp Asn Leu Cys Ser Lys Tyr Pro Glu Leu  
195 200 205  
Leu Gly Ile Gln His Ala Pro Ile Thr Arg Lys Thr Asp Leu Glu Ser  
210 215 220  
Ser Pro Val Trp Trp Phe Ser Pro Pro Lys Thr Cys Val Leu Met Glu  
225 230 235 240

652

35 40 45  
Ala Ile Ala Gly Ser Leu Val Ser Thr Pro Val Ile Gly Trp Ser Leu  
50 55 60  
Tyr Thr Leu Lys Thr Thr Gly Cys Gly Leu Pro Gly Pro Ala Gly  
65 70 75 80  
Leu Ile Gly Ala Leu Glu Gly Val Ser Tyr Leu Val Val Gly Ile  
85 90 95  
Val Gly Trp Ser Leu Tyr Thr Lys Thr Lys Thr Gly Ser Gly Leu Pro  
100 110  
Asn Gly Pro Phe Gly Leu Leu Gly Ala Val Glu Gly Leu Ser Tyr Leu  
115 120 125  
Ser Val Leu Ala Ile Leu Val Val Phe Gly Ile Glu Phe Leu Asp Asn  
130 135 140  
Gly Ser Val Pro Gly Pro Leu Pro Ser Asp Glu Cys Phe Gly  
145 150 155  
(2) INFORMATION FOR SEQ ID NO:133:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 330 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(E) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..330  
(D) OTHER INFORMATION: / Ceres Seq. ID 2033706  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:  
atgggccaca gaggcagcag agagagagat aagaatggaga agcactctgg aaatgcgtac  
agggggatcc caggagctgc acagcagat gcagcagtt tctgaarat agcattgct  
ctgcgcgctt tctgccttct tactcgggtt aaactctaca ctctcaacag ctctcaacag  
tctcaatccc atgacagctg gatcgtgtt gagcagcacc actggaaaaa agcagaggca  
aagaagcaac cttagctatt caagtgaat aagataccaca aagaagtcag ggactcgtac  
tattcaact gccacgtcta ctccctatg  
60 120 180 240 300  
(2) INFORMATION FOR SEQ ID NO:134:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 109 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(E) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..109  
(D) OTHER INFORMATION: / Ceres Seq. ID 2033707  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:  
Met Gly His Arg Val Thr Arg Glu Arg Val Lys Met Glu Lys Tyr Phe  
1 5 10 15  
Gly Asn Ala Tyr Arg Gly Asp Pro Gly Val Pro His Ala Asp Ala Asp  
20 25 30  
Arg Phe Val Asn Ile Trp Ile Gly Ser Ala Ala Phe Ser Val Leu Thr  
35 40 45  
Trp Val Asn Pro Tyr Met Trp Glu Leu Ser Asn Glu Phe Asn Tyr His  
50 55 60  
Asp Lys Trp Met Leu Phe Glu Glu Tyr His Trp Lys Lys Ala Arg Ala  
65 70 75 80  
Lys Lys Glu Pro Tyr Glu Phe Lys Trp Asn Lys Ile Pro Lys Glu Val  
85 90 95  
Arg Asp Ser Tyr Tyr Tyr Asn Trp Pro Val Tyr Phe Pro  
100 105  
(2) INFORMATION FOR SEQ ID NO:135:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 98 amino acids

653

(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..98  
(D) OTHER INFORMATION: / Ceres Seq. ID 2033708  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:  
Met Glu Lys Tyr Phe Gly Asn Ala Tyr Arg Gly Asp Pro Gly Val Pro 15  
1 His Ala Asp Ala Asp Arg Phe Val Asn Ile Trp Ile Gly Ser Ala Ala 30  
Phe Ser Val Leu Thr Trp Val Asn Pro Tyr Met Trp Gln Leu Ser Asn 45  
15 Gln Phe Asn Tyr His Asp Lys Trp Met Leu Phe Glu Gln Tyr His Trp 60  
Lys Lys Ala Arg Ala Lys Lys Gln Pro Tyr Glu Phe Lys Trp Asn Lys 75  
65 Ile Pro Lys Glu Val Arg Asp Ser Tyr Tyr Asn Trp Pro Val Tyr 90  
Phe Pro 95

(2) INFORMATION FOR SEQ ID NO:136:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 56 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..56  
(D) OTHER INFORMATION: / Ceres Seq. ID 2033709

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:136:

Met Trp Gln Leu Ser Asn Gln Phe Asn Tyr His Asp Lys Trp Met Leu 15  
1 Phe Glu Gln Tyr His Trp Lys Lys Ala Arg Ala Lys Lys Gln Pro Tyr 30  
Glu Phe Lys Trp Asn Lys Ile Pro Lys Glu Val Arg Asp Ser Tyr Tyr 45  
Tyr Asn Trp Pro Val Tyr Phe Pro 50  
55

(2) INFORMATION FOR SEQ ID NO:137:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 418 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: DNA (genomic)  
(ix) FEATURE:  
(A) NAME/KEY: -  
(B) LOCATION: 1..418  
(D) OTHER INFORMATION: / Ceres Seq. ID 2043118

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:

ataacaatat caaaagataa aaacagatt tgcctttata agatcaaga aggtaaaaa 60  
agaattgga aatggagtg cgtcgtatt gataatgat gtagcatgtg ttgcgtcac 120  
aatagaaga caagaagaaa gtgggggac taattgtttt gccaatgtt ctgccttgc 180  
tcgacgtac gacggcagt gcatataaa ctgtaaaac gaacgggtg gtccaaaggc 240  
tcctttat cgtctcaga gtcacacgc atgacatggc atcatcaatg gaagttcgc 300  
gaagatagg atgataaaa gagatttga gcggatttt gatcatctt actgtttta 360  
acttcacgtt ttatattata agattttat taaagatgc aataaacga gaattgt

654

(2) INFORMATION FOR SEQ ID NO:138:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 69 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..69  
(D) OTHER INFORMATION: / Ceres Seq. ID 2043119

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:138:

Met Ala Lys Trp Ser Ala Ile Val Leu Ile Met Met Val Met Ile Val 15  
1 Ala Val Thr Ile Glu Ala Gln Glu Ser Gly Gly Thr Ile Cys Phe 30  
Arg Gln Cys Ser Gln Pro Cys Arg Ser Asp Asp Gly Ser Cys Tyr Glu 45  
35 Asn Cys Lys Ile Glu Cys Gly Gly Pro Lys Pro Pro Leu Ser Arg Leu 60  
Arg Ser Ser His Ala 65

(2) INFORMATION FOR SEQ ID NO:139:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 62 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..62  
(D) OTHER INFORMATION: / Ceres Seq. ID 2043120

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:139:

Ile Thr Ile Ser Lys Asp Lys Asn Arg Val Cys Phe Tyr Lys Ile Lys 15  
1 Glu Gly Lys Lys Asn Gly Lys Met Glu Cys Asp Arg Ile Asp Asn 30  
Asp Gly Asp Asp Cys Cys Gly His Asn Arg Ser Thr Arg Arg Lys Trp 45  
Trp Asp Tyr Leu Phe Ser Pro Met Phe Ser Ala Leu Ser Gln 60  
55

(2) INFORMATION FOR SEQ ID NO:140:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 59 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  
(ii) MOLECULE TYPE: peptide  
(ix) FEATURE:  
(A) NAME/KEY: peptide  
(B) LOCATION: 1..59  
(D) OTHER INFORMATION: / Ceres Seq. ID 2043121

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:140:

Met Met Val Met Ile Val Ala Val Thr Ile Glu Ala Gln Glu Glu Ser 15  
1 Gly Gly Thr Ile Cys Phe Arg Gln Cys Ser Gln Pro Cys Arg Ser Asp 30  
20 Asp Gly Ser Cys Tyr Glu Asn Cys Lys Ile Glu Cys Gly Gly Pro Lys 45  
35 Pro Pro Leu Ser Arg Leu Arg Ser Ser His Ala 55  
59



655

(2) INFORMATION FOR SEQ ID NO:141:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 838 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..838

(D) OTHER INFORMATION: / Ceres Seq. ID 2047214

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:141:

15 acacgaacg atccctctc tcgcttctc tcataatat ctcactcat gctcttagt  
 tgccttcca caacgaatt tcattcttc cttacagta aacgaatctc tctctcaaa  
 gttttatt cttctcgga tgcctcaaga agatttact gctgttgta ctaacgtgc  
 tggctccgtg gagacgacac ttgtctcac tgaattcaag cagatttgc tcttgaagc  
 tcaaaagtc ggtgacgctg ttaatttcta caaatctgct ttgtgtcga ctagctctg  
 tcaattctt taccctaacg gtaacttga ccaagagctt cctcatgttc tctctctga  
 gcttaattc gctggctctt cctctgtgt ttgcagctt tctctctcc ctggttttc  
 tactcgaaa tcgaaaggtt cggagtgac ttctcttc ggaactaag atcgaagc  
 cgcgttgc aagctgttg accctggagc tctgaagtg gagctacgg agcagaagt  
 gaaactgga tcaaaagaa aagttacgga tcttttgt gtaacttga tcttcgga  
 gaagaacac gtatcacgc acgaagaaca agagtttag atctctgc cgaactaat  
 tctctcgcg ttctgaaca aaaaaacta tctattacc taggttag ctctatttc  
 20 gctcttgg gaacccgga atcaactaac cggactcagt agcraatga cttcggatt  
 ctaattcac ccttttgc aactctaat cttgtgaaa ttgatatta atctcggg  
 (2) INFORMATION FOR SEQ ID NO:142:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 166 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..166  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2047215  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:142:  
 Met Ala Gln Glu Asp Val Thr Ala Val Thr Asn Gly Ala Gly Pro  
 1 Val Glu Thr His Leu Val Phe Thr Glu Phe Lys Gln Met Leu Val  
 Glu Ala Gln Lys Val Gly Asp Ala Val Thr Phe Tyr Lys Ser Ala Phe  
 35 Gly Ala Ile Glu Ser Gly His Ser Leu Tyr Pro Lys Arg Lys Leu Asp  
 Gln Glu Leu Pro His Val Leu Ser Ser Glu Leu Asn Leu Ala Gly Ser  
 45 Ser Phe Val Val Cys Asp Val Ser Ser Leu Pro Gly Phe Ser Thr Ala  
 Lys Ser Glu Gly Ser Gly Val Thr Phe Leu Leu Gly Thr Lys Asp Ala  
 Glu Ala Val Ala Lys Ala Val Asp Ala Gly Ala Val Lys Val Glu  
 55 Val Thr Glu Ala Glu Val Glu Leu Gly Phe Lys Gly Lys Val Thr Asp  
 Pro Phe Gly Val Thr Trp Ile Phe Ala Glu Lys Lys Thr Val Ile Thr  
 145 Asp Glu Asn Lys Glu Val  
 165  
 (2) INFORMATION FOR SEQ ID NO:143:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 138 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..138  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2047216  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:  
 Met Leu Leu Val Glu Ala Gln Lys Val Gly Asp Ala Val Thr Phe Tyr  
 1 Lys Ser Ala Phe Gly Ala Ile Glu Ser Gly His Ser Leu Tyr Pro Lys  
 10 Arg Lys Leu Asp Gln Glu Leu Pro His Val Leu Ser Ser Glu Leu Asn  
 Leu Ala Gly Ser Ser Phe Val Val Cys Asp Val Ser Ser Leu Pro Gly  
 15 Phe Ser Thr Ala Lys Ser Glu Gly Ser Gly Val Thr Phe Leu Leu Gly  
 Thr Lys Asp Ala Glu Ala Val Ala Lys Ala Val Asp Ala Gly Ala  
 20 Val Lys Val Glu Val Thr Glu Ala Glu Val Glu Leu Gly Phe Lys Gly  
 Lys Val Thr Asp Pro Phe Gly Val Thr Trp Ile Phe Ala Glu Lys Lys  
 25 Thr Val Ile Thr Asp Glu Asn Lys Glu Val  
 130  
 (2) INFORMATION FOR SEQ ID NO:144:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 98 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..98  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2047217  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:  
 Met Leu Leu Leu Leu Leu Leu Thr Val Leu Val Arg Trp Arg Arg Ile  
 1 Leu Ser Ser Leu Ser Ser Arg Cys Ser Leu Lys Leu Lys Arg  
 Ser Val Thr Leu Leu Leu Ser Thr Asn Leu Leu Val Arg Ser Ser  
 45 Leu Val Ile Leu Phe Thr Leu Ser Val Ser Leu Thr Lys Ser Phe Leu  
 Met Phe Ser Leu Leu Ser Leu Ile Ser Leu Ala Leu Pro Ser Leu Phe  
 50 Ala Thr Phe Pro Leu Ser Leu Val Phe Leu Leu Arg Asn Arg Lys Val  
 Arg Glu  
 90  
 (2) INFORMATION FOR SEQ ID NO:145:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 682 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:

656

(A) LENGTH: 138 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..138

(D) OTHER INFORMATION: / Ceres Seq. ID 2047216

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:

10 Met Leu Leu Val Glu Ala Gln Lys Val Gly Asp Ala Val Thr Phe Tyr  
 Lys Ser Ala Phe Gly Ala Ile Glu Ser Gly His Ser Leu Tyr Pro Lys  
 15 Arg Lys Leu Asp Gln Glu Leu Pro His Val Leu Ser Ser Glu Leu Asn  
 Leu Ala Gly Ser Ser Phe Val Val Cys Asp Val Ser Ser Leu Pro Gly  
 Phe Ser Thr Ala Lys Ser Glu Gly Ser Gly Val Thr Phe Leu Leu Gly  
 Thr Lys Asp Ala Glu Ala Val Ala Lys Ala Val Asp Ala Gly Ala  
 Val Lys Val Glu Val Thr Glu Ala Glu Val Glu Leu Gly Phe Lys Gly  
 Lys Val Thr Asp Pro Phe Gly Val Thr Trp Ile Phe Ala Glu Lys Lys  
 Thr Val Ile Thr Asp Glu Asn Lys Glu Val  
 130  
 (2) INFORMATION FOR SEQ ID NO:144:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 98 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: peptide  
 (ix) FEATURE:  
 (A) NAME/KEY: peptide  
 (B) LOCATION: 1..98  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2047217  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:  
 Met Leu Leu Leu Leu Leu Thr Val Leu Val Arg Trp Arg Arg Ile  
 1 Leu Ser Ser Leu Ser Ser Arg Cys Ser Leu Lys Leu Lys Arg  
 Ser Val Thr Leu Leu Leu Ser Thr Asn Leu Leu Val Arg Ser Ser  
 45 Leu Val Ile Leu Phe Thr Leu Ser Val Ser Leu Thr Lys Ser Phe Leu  
 Met Phe Ser Leu Leu Ser Leu Ile Ser Leu Ala Leu Pro Ser Leu Phe  
 50 Ala Thr Phe Pro Leu Ser Leu Val Phe Leu Leu Arg Asn Arg Lys Val  
 Arg Glu  
 90  
 (2) INFORMATION FOR SEQ ID NO:145:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 682 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:

657

(A) NAME/KEY: -  
 (B) LOCATION: 1..682  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2047438  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:  
 5 atgattggtg gttttctcgc cagatgaataa ttcaacagat cagagagaga aaaactctct 60  
 ctctctctc cttccacagt cgcactcgc cttacagta cattttgtt agcaattctg 120  
 tcttttagg attcaatgt ttgcgcaga ggataataa tcttccatg ctgagcttc 180  
 gtccaaatt ttcttatac ttgtggttc ggtgtagct gatgaacat ctgaggtca 240  
 cgggtaaca cgaataagac ttgatgtga ttggatata gtagaagag agttgagatt 300  
 gtagtgaa ggggtgcga agatcctga tcttagcac aactctgaa ccaacacca 360  
 aatgttgt gatatttt gtagctga cctctctga gctagtga gtttcaattg 420  
 catgaactg ggggacaaa ttgtcagc aggtttgt cctcatctg aaaaatgat 480  
 gggaaagga aagaagctc ggcacagc accaagac acaagctg cacagaacg 540  
 gaatgcaga cgcagccca atccagata ttctctat ccaattctg ctagtgaag 600  
 aagaactg aaaggaact ga  
 (2) INFORMATION FOR SEQ ID NO:146:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 181 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ix) FEATURE:  
 (i) NAME/KEY: peptide  
 (B) LOCATION: 1..181  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2047439  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:146:  
 5 Met Ser Gly Ala Glu Asp Asn Lys Ser Ser His Ala Gln Leu Ser Ser 15  
 10 Gln Ile Phe Leu Asp Leu Val Asp Ser Val Ile Ala Asp Val Ala Ser 20  
 25 Glu Cys His Arg Val Ala Arg Leu Gly Leu Asp Arg Asp Leu Ile 30  
 35 Val Glu Glu Glu Leu Arg Leu Ser Val Glu Ala Arg Ala Lys Ile Ala 40  
 45 Asp Pro Ser Asn Asn Leu Glu Thr Asn Thr Lys Tyr Val Val Asp Ile 50  
 55 Phe Gly Gln Thr His Pro Pro Val Ala Ser Glu Val Phe Asn Cys Met 60  
 65 Asn Cys Gly Arg Gln Ile Val Ala Gly Arg Phe Ala Pro His Leu Glu 70  
 75 Lys Cys Met Gly Lys Gly Arg Lys Ala Arg Ala Lys Thr Thr Arg Ser 80  
 85 Thr Thr Ala Ala Gln Asn Arg Asn Ala Arg Arg Ser Pro Asn Pro Arg 90  
 95 Tyr Ser Pro Tyr Pro Asn Ser Ala Ser Glu Asn Gln Leu Ala Ser Gly 100  
 105 Ser Pro Gly Val Ala Gly Glu Asp Cys Ser Asn Phe Thr Val Arg Glu 110  
 115 Asn Val Lys Gly Asp 120  
 125  
 130  
 135  
 140  
 145  
 150  
 155  
 160  
 165  
 170  
 175  
 180  
 (2) INFORMATION FOR SEQ ID NO:147:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 501 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ix) FEATURE:  
 (i) NAME/KEY: -  
 (B) LOCATION: 1..501

658

(D) OTHER INFORMATION: / Ceres Seq. ID 2049056  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:147:  
 5 aaacaaaggt gtaagaaga aggtacaaca tattaattga ctaaaaaatga aggcattggt 60  
 gataacttg ttggtgatt gttcgtcgt ggtggtggag caatcagagg ctccaaagg 120  
 tgaagaatg ttaaatccag gctgcttga ccggtgctg ggtcctaate cccagcggg 180  
 atgcatctt ccaattccc accaaaacc tgcgtccct gtcccaatt atagtcgtg 240  
 ttgtgtaga atctacagt gcagagaga tgcctaggt taactcaca tgacgagac 300  
 gttcccttat attgactct ctctacatg agacttgcaa tgatatata tattcatata 360  
 tatcacatt aagcgtgata tgataatcc agattgttaa tataaaatc ttactgatt 420  
 caatcaca cacaagatcc tgaactgaa ataatttgt ccatgtttg ttgtataat 480  
 aatgttga cttctattat t  
 (2) INFORMATION FOR SEQ ID NO:148:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 76 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ix) FEATURE:  
 (i) NAME/KEY: peptide  
 (B) LOCATION: 1..76  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2049057  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:148:  
 5 Met Lys Ala Trp Met Ile Ile Leu Leu Val Ile Cys Val Ala Val Val 15  
 10 Val Gln Gln Ser Glu Ala Arg Lys Arg Lys Tyr Leu Asn Pro Gly 20  
 25 Val Leu Asp Arg Cys Arg Gly Pro Asn Pro Pro Ala Gly Cys His Pro 30  
 35 His Asn Ser His Lys Lys Pro Arg Val Pro Val His Asn Tyr Ser Arg 40  
 45 Gly Cys Ser Arg Ile Thr Arg Cys Arg Arg Asp Ala 50  
 55  
 60  
 65  
 70  
 75  
 (2) INFORMATION FOR SEQ ID NO:149:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 72 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear  
 (ix) FEATURE:  
 (i) NAME/KEY: peptide  
 (B) LOCATION: 1..72  
 (D) OTHER INFORMATION: / Ceres Seq. ID 2049058  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:149:  
 5 Met Ile Ile Leu Leu Val Ile Cys Val Ala Val Val Val Gln Gln Ser 15  
 10 Glu Ala Arg Lys Gly Arg Lys Tyr Leu Asn Pro Gly Val Leu Asp Arg 20  
 25 Cys Arg Gly Pro Asn Pro Pro Ala Gly Cys His Pro His Asn Ser His 30  
 35 His Lys Pro Arg Val Pro Val His Asn Tyr Ser Arg Gly Cys Ser Arg 40  
 45 Ile Thr Arg Cys Arg Arg Asp Ala 50  
 55  
 60  
 65  
 70  
 75  
 (2) INFORMATION FOR SEQ ID NO:150:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1007 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ix) FEATURE:  
 (i) NAME/KEY: DNA (genomic)

659

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..1007

(D) OTHER INFORMATION: / Ceres Seq. ID 2050386

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:150:

5 ccaaatcttc ccaaggatgc tcttggat cctccgatt tagccttagg gttcgatgag 60  
 cagatagcta ttcatacgcc aagaataac taatagaaga ctgtgatgag 120  
 acgagtgtgt ctttgtact taccagcta atgaacgtg tctctccc taactccc 180  
 gaacacatla tctttctgc tttgtcgc cctttcttc attatagc aacttgctt 240  
 aaactggat gcaatttagc taccataac tgaataatc ctgtgattt cttgacatg 300  
 ctcatggta agtgttcaga tgggtacaa agtcagaca atgtgagtc agttgcgaaa 360  
 ctattcggg agatacaga aaccttcga agctacaga atgtaacag tgaacata 420  
 actgtatgg tggatgact gttctgtg gaattgcta ctaccggag caactcagat 480  
 cactatgag actcttgca tttctgac acataagt ctgaagcaa ttgtcattg 540  
 gtcatactc atcataga tataacgc agtatgga gacttcatt ttgtacag 600  
 atggtatgc ttccagatg ttgtataag gtagagcctt tagctctgg tttagcaat 660  
 gatcacatg gccaatgac ttcttgac aaaggataa gcaactcag taguagaac 720  
 tccggaaac agttgcagaa ttccaattc agtatcaa gaaatgtat cgactatttc 780  
 tatcttggt gcagaactg agattagac acctctgc gattagcat ggtatgaga 840  
 tctcttcta gacagtttg ttatttgt ttaggttaa gagagacat ctttaasca 900  
 ttttactt agtgcctt ctgttgta cttgatgact tagctaat tgaatgcttg 960  
 agacacaga acctgatag atggaaatt agtaacattc tatattc

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 266 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..266

(D) OTHER INFORMATION: / Ceres Seq. ID 2050387

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:151:

35 Pro Asn Phe Ser Met Asp Arg Ser Leu Asn Leu Leu Asp Leu Ala Leu 15  
 1 Gly Phe Asp Glu Gln Leu Ala Ile Pro Ser Pro Leu Asn Gly Lys Val 20  
 Ile Leu Ile Glu Asp Cys Val Glu Thr Ser Gly Ser Phe Val Leu His 25  
 Gln Leu Met Lys Arg Val Leu Ser Ser Asn Ser Ser Asp Ala Leu Ile 30  
 Phe Leu Ala Phe Ala Arg Pro Phe Ser His Tyr Asp Arg Ile Leu Arg 35  
 Lys Leu Gly Cys Asn Leu Ala Thr His Lys Ser Ser Asn Arg Leu Val 40  
 Phe Phe Asp Met Leu Met Val Lys Cys Ser Asp Gly Asp Gln Met Glu 45  
 Asp Asn Val Ser Ala Val Ala Lys Leu Phe Arg Glu Ile Gln Glu Thr 50  
 Val Arg Lys Leu Gln Ser Val Thr Ser Gly Asn Ile Thr Val Met Val 55  
 Asp Asp Met Ser Leu Leu Glu Ile Ala Thr Thr Gly Ser Asn Ser Asp 60  
 His Val Leu Asp Phe Leu His Tyr Cys His Thr Leu Ser Ser Glu Ser 65  
 Asn Cys Ser Leu Val Ile Leu Asn His Glu Asp Ile Tyr Ala Ser Met 70  
 Glu Arg Pro Ala Phe Leu Leu Gln Met Val Cys Leu Ala Asp Val 75  
 Ile Lys Ala Glu Pro Leu Ala Ser Gly Leu Ala Asn Asp Val His Gly 80  
 210

660

Gln Leu Thr Val Leu Asn Lys Gly Ile Ser Asn Ser Gly Arg Gly Ser 225  
 Ser Arg Asn Lys Leu Gln Asn Phe Gln Phe Arg Ile Lys Glu Asn Gly 240  
 245  
 250  
 255

5 Ile Asp Tyr Phe Thr Pro Gly Cys Arg Ser 260

(2) INFORMATION FOR SEQ ID NO:152:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 262 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..262

(D) OTHER INFORMATION: / Ceres Seq. ID 2050388

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:152:

Met Asp Arg Ser Leu Asn Leu Asp Leu Leu Ala Leu Gly Phe Asp Glu 1  
 5  
 10  
 15  
 20 Gln Leu Ala Ile Pro Ser Pro Leu Asn Gly Lys Val Ile Leu Ile Glu  
 Asp Cys Val Glu Thr Ser Gly Ser Phe Val Leu His Gln Leu Met Lys 35  
 Arg Val Leu Ser Ser Asn Ser Ser Asp Ala Leu Ile Phe Leu Ala Phe 45  
 Ala Arg Pro Phe Ser His Tyr Asp Arg Ile Leu Arg Lys Leu Gly Cys 50  
 65 Asn Leu Ala Thr His Lys Ser Asn Asn Arg Leu Val Phe Phe Asp Met 75  
 85  
 90  
 95  
 100  
 105  
 110  
 115  
 120  
 125  
 130  
 135  
 140  
 145  
 150  
 155  
 160  
 165  
 170  
 175  
 180  
 185  
 190  
 195  
 200  
 205  
 210  
 215  
 220  
 225  
 230  
 235  
 240  
 245  
 250  
 255  
 Tyr Pro Gly Cys Arg Ser 260

(2) INFORMATION FOR SEQ ID NO:153:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 216 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..216

661

(D) OTHER INFORMATION: / Ceres Seq. ID 2050389

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:153:

Met Lys Arg Val Leu Ser Ser Asn Ser Asp Ala Leu Ile Phe Leu

1

Ala Phe Ala Arg Pro Phe Ser His Tyr Asp Arg Ile Leu Arg Lys Leu

5

Gly Cys Asn Leu Ala Thr His Lys Ser Asn Asn Arg Leu Val Phe Phe

10

Asp Met Leu Met Val Lys Cys Ser Asp Gly Asp Gln Met Glu Asp Asn

15

Val Ser Ala Val Ala Lys Leu Phe Arg Glu Ile Gln Glu Thr Val Arg

20

Lys Leu Gln Ser Val Thr Ser Gly Asn Ile Thr Val Met Val Asp Asp

25

Met Ser Leu Leu Glu Ile Ala Thr Thr Gly Ser Asn Ser Asp His Val

30

Leu Asp Phe Leu His Tyr Cys His Thr Leu Ser Ser Glu Ser Asn Cys

35

Ser Leu Val Ile Leu Asn His Glu Asp Ile Tyr Ala Ser Met Glu Arg

40

Pro Ala Phe Leu Leu Gln Met Val Cys Leu Ala Asp Val Val Ile Lys

45

Ala Glu Pro Leu Ala Ser Gly Leu Ala Asn Asp Val His Gly Gln Leu

50

Thr Val Leu Asn Lys Gly Ile Ser Asn Ser Gly Arg Gly Ser Ser Arg

55

Asn Lys Leu Gln Asn Phe Gln Phe Arg Ile Lys Glu Asn Gly Ile Asp

60

Tyr Phe Tyr Pro Gly Cys Arg Ser

65

(2) INFORMATION FOR SEQ ID NO:154:

(A) LENGTH: 634 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..634

(D) OTHER INFORMATION: / Ceres Seq. ID 2053353

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:154:

gaataaaat cttattcttg tgttggtt gttattctt aaaaaataa agagatcag

45

tttgggagaa agcagcaag aagaataatg gcagcaaat ttgcaaac atcgaggtg

50

atagacctg gagatcacc catcaccacc aacctctt cttatcctt ttaaaaca

55

acattaaagg caagaaccc ttgagactc gcggtgcat cggagagaa agtcactgc

60

ttgagagga ctaggtttag gagatattg aacgttagt gatttggtt gtcgagtg

65

ctagctccgt ctagctacc agcagataat ggaagagcc tgaagggtt cttctgat

70

agcaccgaa ctgagctgc tcactcccg actctccag cttcactc acagttcgg

75

ttgggttgg ttacgtggc ttacggctc ttccctgcg tcaatttcg acaatcgg

80

ttcaaggcca ggaactaga ttacttcaa ggaactat tctgtttc ttgtacta

85

tagctctt cttgttatg tgaatatta tcaatcaaa caaagcttc ttgtattt

90

gatgttcat tctatctc gcaagttt tctt

95

(2) INFORMATION FOR SEQ ID NO:155:

(A) LENGTH: 140 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..130

(D) OTHER INFORMATION: / Ceres Seq. ID 2055694

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:155:

Met Lys Lys Pro Leu Ser Leu Ser Gln Phe Thr Ser Pro Pro Ser Met

55

Ala Trp Arg Asn Ala Gly Ser Ala Ala Arg Ser Phe Val Ser Ala Thr

60

Ala Arg Ser Pro Ser Leu Arg Ser Pro Thr Thr Ala Leu Pro Arg Leu

65

Arg Pro Pro Gln Ser Ser Leu Ser Arg Arg Phe Thr Phe Ser Ser

70

Pro Ser Arg Asn Leu Gly Ala Leu Gly Cys Thr Gln Ser Phe Leu Pro

75

Met Lys Lys Pro Leu Ser Leu Ser Gln Phe Thr Ser Pro Pro Ser Met

80

Ala Trp Arg Asn Ala Gly Ser Ala Ala Arg Ser Phe Val Ser Ala Thr

662

(D) OTHER INFORMATION: / Ceres Seq. ID 2053354

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:156:

Met Ala Ala Thr Phe Ala Thr Pro Ser Thr Val Ile Gly Leu Gly Gly

5

Ser Ser Ile Thr Thr Lys Pro Phe Ser Ser Ser Phe Leu Lys Pro Thr

10

Leu Ser Ala Lys Asn Pro Leu Arg Leu Ala Gly Ala Ser Gly Arg

15

Val Thr Cys Phe Glu Arg Asn Trp Leu Arg Arg Asp Leu Asn Val Val

20

Gly Phe Gly Leu Ile Gly Trp Leu Ala Pro Ser Ile Pro Ala Ile

25

Asn Gly Lys Ser Leu Thr Gly Leu Phe Phe Asp Ser Ile Gly Thr Glu

30

Leu Ala His Phe Pro Thr Pro Pro Ala Leu Thr Ser Gln Phe Thr Gly

35

Trp Leu Val Thr Trp His Leu Gly Leu Phe Leu Cys Leu Thr Phe Gly

40

Gln Ile Gly Phe Lys Gly Arg Thr Glu Asp Tyr Phe

45

(2) INFORMATION FOR SEQ ID NO:157:

(A) LENGTH: 393 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..393

(D) OTHER INFORMATION: / Ceres Seq. ID 2055693

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:157:

atgcaaaaag cctctctct cttcaaatc acctctctc ttcaatgac ttggcgcaac

50

ccacacgg cgattcttcg cctcgtctt cttcaatct cttacttag cctcgtctt

55

acctctat cactctccag gaattagga gaattggtt gcacacagc gttctgctt

60

ctgtacagtg ttgggtctac ttcaaatc acctctacc ttaactgaa ttggagact

65

ctggcgagc tgcctacgg gaattgaaa agatgggta tgaactcga tttaactgcg

70

ctggcaagag aatgggtcg cactgtccg tga

75

(2) INFORMATION FOR SEQ ID NO:158:

(A) LENGTH: 130 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..130

(D) OTHER INFORMATION: / Ceres Seq. ID 2055694

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:158:

Met Gln Lys Pro Leu Ser Leu Ser Gln Phe Thr Ser Pro Pro Ser Met

80

Ala Trp Arg Asn Ala Gly Ser Ala Ala Arg Ser Phe Val Ser Ala Thr

85

Ala Arg Ser Pro Ser Leu Arg Ser Pro Thr Thr Ala Leu Pro Arg Leu

90

Arg Pro Pro Gln Ser Ser Leu Ser Arg Arg Phe Thr Phe Ser Ser

95

Pro Ser Arg Asn Leu Gly Ala Leu Gly Cys Thr Gln Ser Phe Leu Pro

100

Met Lys Lys Pro Leu Ser Leu Ser Gln Phe Thr Ser Pro Pro Ser Met

105

Ala Trp Arg Asn Ala Gly Ser Ala Ala Arg Ser Phe Val Ser Ala Thr

110

Ala Arg Ser Pro Ser Leu Arg Ser Pro Thr Thr Ala Leu Pro Arg Leu

115

Arg Pro Pro Gln Ser Ser Leu Ser Arg Arg Phe Thr Phe Ser Ser

120

Pro Ser Arg Asn Leu Gly Ala Leu Gly Cys Thr Gln Ser Phe Leu Pro

125

Met Lys Lys Pro Leu Ser Leu Ser Gln Phe Thr Ser Pro Pro Ser Met

130

Ala Trp Arg Asn Ala Gly Ser Ala Ala Arg Ser Phe Val Ser Ala Thr

135

Ala Arg Ser Pro Ser Leu Arg Ser Pro Thr Thr Ala Leu Pro Arg Leu

140

Arg Pro Pro Gln Ser Ser Leu Ser Arg Arg Phe Thr Phe Ser Ser

145

Pro Ser Arg Asn Leu Gly Ala Leu Gly Cys Thr Gln Ser Phe Leu Pro

150

Met Lys Lys Pro Leu Ser Leu Ser Gln Phe Thr Ser Pro Pro Ser Met

663

Leu Tyr Ser Val Val Ala Thr Ser Gln Leu Thr Ser His Leu Asn Val  
 85 90  
 Asn Leu Arg Ala Phe Cys Glu Leu Ser Asn Gly Asn Trp Lys Arg Trp  
 100 105  
 Val Met His Ser Asp Phe Thr Ser Leu Ala Arg Glu Trp Cys Ile  
 115 120  
 Val Pro

(2) INFORMATION FOR SEQ ID NO:158:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 115 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..115

(D) OTHER INFORMATION: / Ceres Seq. ID 2055695

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:158:

Met Ala Trp Arg Asn Ala Gly Ser Ala Ala Arg Ser Phe Val Ser Ala  
 1 5 10 15  
 Thr Ala Arg Ser Pro Ser Leu Arg Ser Pro Thr Thr Ala Leu Pro Arg

20

Leu Arg Pro Pro Gln Ser Ser Leu Pro Ser Arg Arg Phe Thr Phe Ser  
 20 25 30 35 40 45

Ser Pro Ser Arg Asn Leu Gly Ala Leu Gly Cys Thr Gln Ser Phe Leu  
 50 55 60 65 65 70 75 80

Pro Leu Tyr Ser Val Val Ala Thr Ser Gln Leu Thr Ser His Leu Asn  
 70 75 80 85 90 95

Val Asn Leu Arg Ala Phe Cys Glu Leu Ser Asn Gly Asn Trp Lys Arg  
 90 95 100 105 110

Trp Val Met His Ser Asp Phe Thr Ser Leu Ala Arg Glu Trp Cys Cys  
 100 105 110 115

Ile Val Pro

(2) INFORMATION FOR SEQ ID NO:159:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 68 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..68

(D) OTHER INFORMATION: / Ceres Seq. ID 2055696

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:159:

Cys Lys Asn Leu Ser Leu Ser Leu Asn Ser Leu Leu Leu Gln Trp  
 1 5 10 15

Leu Gly Ala Thr Gln Asp Leu Pro Leu Val Leu Ser Ser Pro Pro  
 20 25 30 35 40 45

Gln Asp His His Leu Ser Val Leu Pro Pro Arg Arg Phe Leu Ala Ser  
 35 40 45 50 55 60

Val Leu Leu Asn Pro Pro Tyr Leu Ala Val Ala Ser Pro Ser His His  
 50 55 60 65

Leu Pro Gly Ile

(2) INFORMATION FOR SEQ ID NO:160:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 394 base pairs

(B) TYPE: nucleic acid

664

(C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..394

(D) OTHER INFORMATION: / Ceres Seq. ID 2056405

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:160:

tcaatagctt atctagttct tgaagtcca acgatacaatg gcaccatccg ctgcaatgct  
 catactcca catctctag ttacgtacaa agctaaagaat cagttctgt catcgcgctc  
 gtgttaag tgcacgtg tcttggttt tcttggcct tggaaagcat tagacaatga  
 ggatattcc gcggtgttc tggcgagct ctctgggcat cggctacta tcgaagaacg  
 ctccaaga gctttgaac aaagtgttg gtaattgtg caattgttg ttgggttat  
 tgaatttt tagcttaaaa tagctttat aaattgaaa tggaaagtact gtaaatgta  
 aaattgact atataaatt aaagtacat tatg

(2) INFORMATION FOR SEQ ID NO:161:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 78 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..78

(D) OTHER INFORMATION: / Ceres Seq. ID 2056406

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:161:

Met Ala Pro Ser Ala Ala Met Leu Ile Leu Ser His Pro Leu Val Ser  
 1 5 10 15

His Lys Ala Lys Asn Gln Ser Leu Ser Ser Pro Ser Val Lys Ser  
 20 25 30 35 40 45

Thr Arg Val Phe Gly Phe Leu Trp Pro Trp Lys Ala Leu Asp Asn Glu  
 50 55 60 65

Asp His Ser Ala Val Val Leu Gly Arg Leu Phe Gly Asp Pro Ala Thr  
 70 75

Ile Glu Lys Arg Phe Gln Glu Ala Leu Glu Gln Ser Cys Trp

(2) INFORMATION FOR SEQ ID NO:162:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 72 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..72

(D) OTHER INFORMATION: / Ceres Seq. ID 2056407

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:162:

Met Leu Ile Leu Ser His Pro Leu Val Ser His Lys Ala Lys Asn Gln  
 1 5 10 15

Ser Leu Ser Ser Pro Ser Ser Val Lys Ser Thr Arg Val Phe Gly Phe  
 20 25 30 35 40 45

Leu Trp Pro Trp Lys Ala Leu Asp Asn Glu Asp His Ser Ala Val Val  
 50 55 60 65

Leu Gly Arg Leu Phe Gly Asp Pro Ala Thr Ile Glu Lys Arg Phe Gln  
 70 75

Glu Ala Leu Glu Gln Ser Cys Trp

(2) INFORMATION FOR SEQ ID NO:163:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 48 amino acids

665

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..48

(D) OTHER INFORMATION: / Ceres Seq. ID 2056408

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:163:

10 Met Arg Ile Ile Pro Arg Leu Phe Ser Ala Gly Ser Leu Ala Ile Arg 15

1 Leu Leu Ser Arg Ser Ala Ser Lys Lys Leu Leu Leu Asn Lys Ala Val Gly 20

15 Asn Leu Tyr Asn Val Trp Leu Gly Leu Leu Ile Val Phe Thr Leu Lys 45

(2) INFORMATION FOR SEQ ID NO:164:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 766 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: -

(B) LOCATION: 1..766

(D) OTHER INFORMATION: / Ceres Seq. ID 2065747

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:164:

30 acatgcctt ccttcctgcg ccgcgcgcgc cgcgcgcgt accaccgct cgcgcgcgc 60

tgcgcctcg cgttcgcga cgcatttgt gaagtttct tgagcgcgcg acatgcgc 120

tgaagacgga actttccgcg ttcagcgcgc agagattta tccgcggaaa ggcattgat 180

ttaccctgc ggaattctcg gcttcctctt ttgcacactc gaattcgaag cgcctacttc 240

accaaccgt gaagctgcga aegcttaact ggcagacaat gtacaggaag cgcacaaga 300

aggtatcca tgcgaagcg gtaagaaga ggcgcgcgcg caccagaag ccatactcca 360

ggtaactgt ggggtcttc ttgaagtaa tccagaaga gagcgttcag aaccagagg 420

tccgcgtgc tgcagagaa gcgcctctc tgcagutcaa ggcgcctc aagaagacca 480

aggtatgaga gaaagcgaag aagcgcgcgc tgagcaagtc ccagaagcg cgcacaagg 540

gtgcgttcca gaaggtttcc aagggcccca agttgcgcgc cgggtgtgg aagcctgaa 600

aqaacttagt ctcgtttcc gaatttcag tgcctctta gccaaagcca cttctgtaga 660

accgtgttg aatttcag acttatatt atgcctgaaa aaaaagacc gttcat 720

(2) INFORMATION FOR SEQ ID NO:165:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 161 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..161

(D) OTHER INFORMATION: / Ceres Seq. ID 2065748

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:165:

Met Val Leu Lys Thr Glu Leu Cys Arg Phe Ser Gly Gln Lys Ile Tyr 15

1 Pro Gly Lys Gly Ile Arg Phe Ile Arg Ala Asp Ser Gln Val Phe Leu 30

55 Phe Ala Asn Ser Lys Cys Lys Arg Tyr Phe His Asn Arg Leu Lys Pro 45

Ala Lys Leu Thr Trp Thr Ala Met Tyr Arg Lys Gln His Lys Lys Asp 60

50 Ile His Ala Glu Ala Val Lys Lys Arg Arg Arg Ala Thr Lys Lys Pro 75

65 80

666

Tyr Ser Arg Ser Ile Val Gly Ala Ser Leu Glu Val Ile Gln Lys Lys 95

Arg Ala Glu Lys Pro Glu Val Arg Asp Ala Ala Arg Glu Ala Ala Leu 110

5 Arg Glu Ile Lys Glu Arg Ile Lys Lys Thr Lys Asp Glu Lys Lys Ala 125

Lys Lys Ala Glu Val Ser Lys Ser Gln Lys Thr Gln Thr Lys Gly Ala 140

10 Val Gln Lys Gly Ser Lys Gly Pro Lys Leu Gly Gly Gly Gly Lys 155

145 Arg 150

(2) INFORMATION FOR SEQ ID NO:166:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 106 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: peptide

(B) LOCATION: 1..106

(D) OTHER INFORMATION: / Ceres Seq. ID 2065749

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:166:

25 Met Tyr Arg Lys Gln His Lys Lys Asp Ile His Ala Glu Ala Val Lys 15

1 Lys Arg Arg Arg Ala Thr Lys Lys Lys Pro Tyr Ser Arg Ser Ile Val Gly 30

Ala Ser Leu Glu Val Ile Gln Lys Lys Arg Ala Glu Lys Pro Glu Val 45

30 Arg Asp Ala Ala Arg Glu Ala Ala Leu Arg Glu Ile Lys Glu Arg Ile 60

Lys Lys Thr Lys Asp Glu Lys Lys Ala Lys Lys Ala Glu Val Ser Lys 75

65 Ser Gln Lys Thr Gln Thr Lys Gly Ala Val Gln Lys Gly Ser Lys Lys 95

Pro Lys Leu Gly Gly Gly Gly Gly Gly Lys Arg 100

## CLAIMS

What is claimed is:

1. An isolated nucleic acid molecule comprising a nucleic acid having a nucleotide sequence selected from the group consisting of
  - (i) a nucleotide sequence shown in SEQ TABLES 1 AND 2;
  - (ii) a nucleotide sequence complementary to one shown in SEQ TABLES 1 AND 2;
  - (iii) a nucleotide sequence which is the reverse complement of one shown in SEQ TABLES 1 AND 2;
  - (iv) a nucleotide sequence capable of hybridizing to a nucleotide sequence shown in SEQ TABLES 1 AND 2;
  - (v) a nucleotide sequence capable of hybridizing to a nucleotide sequence complementary to one shown in SEQ TABLES 1 AND 2;
  - (vi) a nucleotide sequence capable of hybridizing to a nucleotide sequence which is the reverse complement of one shown in SEQ TABLES 1 AND 2;
 whereby the hybridization in groups (iv) to (vi) allow said sequences to form a duplex at a temperature from about T<sub>m</sub>-40°C to about T<sub>m</sub>-48° C.
2. An isolated nucleic acid molecule according to claim 1 that comprises a complete open reading frame.
3. An isolated nucleic acid molecule according to claim 1 that functions as a promoter or as a 3' end termination sequence or as a regulator sequence influencing the transcription rate, the transcript stability or RNA translation rate in a host cell.

4. The isolated nucleic acid of claim 3, comprising a sequence selected from the group consisting of a TATA box sequence, a CAAT box sequence, a motif of GCAATCG or a transcription-factor binding sequence.

5. An isolated nucleic acid molecule comprising a nucleic acid sequence that encodes an amino acid sequence selected from the group consisting of the amino acid sequences shown in SEQ TABLES 1 AND 2, a complementary nucleotide sequence to said nucleic acid sequence that encodes said amino acid sequence selected from SEQ TABLES 1 AND 2, a reverse complementary nucleotide sequence to said nucleic acid sequence that encodes said amino acid sequence selected from SEQ TABLES 1 AND 2, and said amino acid sequence selected from SEQ TABLES 1 AND 2, and a nucleotide sequence that will hybridize to said nucleic acid sequence encoding an amino acid selected from SEQ TABLES 1 AND 2 or said complementary sequence under hybridization conditions providing T<sub>m</sub> - 40°C to 48°C.
6. An isolated nucleic acid molecule comprising a nucleic acid sequence that encodes at least a part or a portion or a mutant or a fusion of an amino acid sequence selected from the group consisting of the amino acid sequences shown in SEQ TABLES 1 AND 2, a complementary nucleotide sequence to said nucleic acid sequence that encodes said part or portion or mutant or fusion of an amino acid sequence selected from SEQ TABLES 1 AND 2, a reverse complementary nucleotide sequence to said nucleic acid sequence that encodes said part or portion or mutant or fusion of an amino acid sequence selected from SEQ TABLES 1 AND 2, and a nucleotide sequence that will hybridize to said nucleic acid sequence that encodes said part or portion or mutant or fusion of an amino acid selected from SEQ TABLES 1 AND 2 or said complementary sequence or said reverse complementary sequence under hybridization conditions providing T<sub>m</sub> - 40°C to T<sub>m</sub> - 48°C.

7. An isolated nucleic acid molecule according to claim 1, having a nucleotide sequence selected from the group consisting of the nucleotide sequences shown in SEQ TABLES 1 AND 2, and a complementary nucleotide sequence to said nucleotide sequence selected from SEQ TABLES 1 AND 2.

8. A chimeric DNA construct comprising (a) a promoter sequence capable of driving gene expression in plant cells and operatively linked to (b) a structural gene comprising an DNA molecule according to any one of claims 1, 2, 5 and 7.

9. A chimeric DNA construct comprising (a) a promoter that is a nucleic acid molecule according to claim 3 or 4 operatively linked to (b) a structural gene and, optionally, (c) an associated UTR.

10. An isolated DNA molecule comprising (a) a promoter sequence according to claim 3 or 4 and operatively linked to (b) a structural gene sequence according to any one of claims 1, 2, 5, 6 and 7.

11. A promoter according to claim 10, wherein said promoter is a specific promoter.

12. A promoter according to claim 11, wherein said promoter is a seed-specific promoter, an embryo-specific promoter, an ovule-specific promoter, a tapetum-specific promoter or a root-specific promoter.

13. A host cell transformed with a polynucleotide comprising the isolated nucleic acid molecule according to claim 1.

14. An isolated polypeptide comprising an amino acid sequence encoded by a polynucleotide sequence shown in SEQ TABLES 1 AND 2, or one that is at least 75% identical thereto.

15. An isolated polypeptide according to claim 14, that is at least 85% identical to said amino acid sequence.

16. An isolated polypeptide, according to claim 15, that is at least 90% identical to said amino acid sequence.

17. A polynucleotide comprising a first polynucleotide sequence from SEQ TABLES 1 AND 2 or a fragment thereof, wherein said first sequence is capable of regulating transcription or translation of second a polynucleotide comprising a coding sequence in a host cell or in vitro system wherein the first and second polynucleotides are operatively linked.

18. The first polynucleotide of claim 17, which is capable of regulating transcription or translation in an in vitro system.

19. The first polynucleotide of claim 17, which is capable of regulating transcription or translation in a host cell.

20. The polynucleotide of claim 19, wherein said host cell is a plant cell.

21. A host cell of claim 13, wherein said isolated nucleic acid comprises a promoter and operatively linked structural gene and further wherein said promoter and structural gene are not heterologous to each other and are exogenous to the host cell genome.



671

22. A method of introducing an isolated nucleic acid into a host cell comprising:

(a) providing an isolated nucleic acid of any of claims 1-12;

5 (b) contacting said isolated nucleic acid with said host cell under conditions that permit insertion of said nucleic acid into said host cell.

23. A method of producing a polypeptide of any one of claims 14-16 comprising:

(a) providing a host cell of claim 13;

5 (b) culturing said host cell under conditions that permit transcription and translation of said structural gene to produce a polypeptide; and

(c) isolating said polypeptide.

24. An antibody capable of binding to a polypeptide of any one of claims 14-16.

25. An isolated nucleic acid comprising a promoter of a gene wherein said gene comprises a nucleic acid having the nucleotide sequence of claim 1.

26. An isolated polypeptide comprising an amino acid sequence selected from SEQ TABLES 1 AND 2, or one that is at least 75% identical thereto.

27. The isolated polypeptide according to Claim 26 that is at least 85% identical thereto.

28. The isolated polypeptide according to claim 26 that is at least 90% identical thereto.

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/00466

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12N 15/29; C07K 14/415  
US CL : 536/23.6; 530/370

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.6; 530/370

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Genbank, Swissprot  
search terms: SEQ ID NO:1-10

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	Database Genbank, Accession No. A1997507, CHEN et al. 'Arabidopsis thaliana Gene Expression Microarray', 'sequence listing', 'unpublished', 08 September 1999.	1, 5, 6, 13
X,P	Database Genbank, Accession No. AT117F15, CHOISNE et al. 'direct submission', 'sequence listing', 'unpublished', 09 June 1999.	1, 5, 6, 13
X,P	Database Genbank, Accession No. T06680, QUETIER et al. 'Hypothetical protein T17F15.100-Arabidopsis thaliana', 'sequence listing', 'unpublished', 23 April 1999.	14, 15, 16, 26, 27, 28

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T
'A' document defining the general state of the art which is not considered to be of particular relevance	have documents published after the international filing date or priority date and which are not considered to be of particular relevance to the principles or theory underlying the invention
'B' earlier documents published on or after the international filing date	documents of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken into account
'L' document which may affect novelty (as priority claim) or which is a document which may affect the inventive step of the invention (as special reasons (as specified))	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is taken into account
'O' document referring to an oral disclosure, use, exhibition or other means	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is taken into account
'P' document published prior to the international filing date but later than the priority date claimed	document number of the same patent family

Date of the actual completion of the international search

02 JUNE 2000

Date of mailing of the international search report

17 JUL 2000

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Preliminary No. (703) 305-3230

Authorized officer

JOHN S. BRUSCA

Telephone No. (703) 300-0196

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/00466

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database Genbank, Accession No. H36097, NEWMAN et al. 'Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones', 'sequence listing', Plant Physiol. (1994), 106, pages 1241-1255, 30 December 1997.	1, 5, 6, 13
X,P	Database Genbank, Accession No. A1999527, CHEN et al. 'Arabidopsis thaliana Gene Expression Microarray', 'sequence listing', 'unpublished', 08 September 1999.	1, 5, 6, 13
X,P	Database Genbank, Accession No. T04228, BEVAN et al. 'untitled', 'sequence listing', 'unpublished', 30 April 1999.	14, 15, 26, 27

Form PCT/ISA/210 (continuation of second sheet) (July 1998)\*

Form PCT/ISA/210 (second sheet) (July 1998)\*

671

22. A method of introducing an isolated nucleic acid into a host cell comprising:

- (a) providing an isolated nucleic acid of any of claims 1-12;
- (b) contacting said isolated nucleic acid with said host cell under conditions that permit insertion of said nucleic acid into said host cell.

5

23. A method of producing a polypeptide of any one of claims 14-16 comprising:

- (a) providing a host cell of claim 13;
- (b) culturing said host cell under conditions that permit transcription and translation of said structural gene to produce a polypeptide; and
- (c) isolating said polypeptide.

5

24. An antibody capable of binding to a polypeptide of any one of claims 14-16.

25. An isolated nucleic acid comprising a promoter of a gene wherein said gene comprises a nucleic acid having the nucleotide sequence of claim 1.

26. An isolated polypeptide comprising an amino acid sequence selected from SEQ TABLES 1 AND 2, or one that is at least 75% identical thereto.

27. The isolated polypeptide according to Claim 26 that is at least 85% identical thereto.

28. The isolated polypeptide according to claim 26 that is at least 90% identical thereto.

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
13 July 2000 (13.07.2000)

(10) International Publication Number  
WO 00/40695 A3

PCT

(51) International Patent Classification: C12N 15/29 (74) Agents: STEWART, Raymond, C. et al.; Birch, Stewart, Kolsch & Birch, LLP, P.O. Box 747, Falls Church, VA 22040-0747 (US).

(21) International Application Number: PCT/US00/00466

(22) International Filing Date: 7 January 2000 (07.01.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data: 8 January 1999 (08.01.1999) US 60/115,293

(71) Applicant (for all designated States except US): CERES, INC. (US/US); 3007 Malibu Canyon Road, Malibu, CA 90265 (US).

(72) Inventors; and  
(75) Inventors/Applicants (for US only): ALEXANDROV, Nikolai (RU/US); 1404 Oak Trail Street, Thousand Oaks, CA 91320 (US); BROVER, Vyacheslav (RU/US); 5916 N. Las Virgenes Road #590, Calabasas, CA 91320 (US); CHEN, Xianfeng (CN/US); 12333 Wild Turkey Court, #B, Creve Coeur, MO 63141 (US); SUBRAMANIAN, Gopalakrishnan (IN/US); 4205 Peach Slope Road, Moorpark, CA 93021 (US); TROUKHAN, Maxim, E. (RU/US); 1671 Amberwood Drive #2, South Pasadena, CA 91030 (US); ZHENG, Liancheng (CN/US); 19212 Circle Gate Drive #201, Germantown, MD 20874 (US).

Published:

— With international search report.

(88) Date of publication of the international search report:  
30 November 2000

For two-letter codes and other abbreviations, refer to the "Guide-  
once Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.

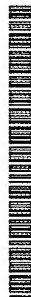
WO 00/40695 A3

(54) Title: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES ENCODED THEREBY

(57) Abstract: The present invention provides DNA molecules that constitute portions of the genome of a plant, and polypeptides encoded thereby. The DNA molecules are useful for expressing a gene product, either as a promoter or as a structural gene or as an UTR or as a 3' termination sequence and are also useful in controlling expression of the target gene or as tools for genetic mapping or identification of a particular individual plant or for clustering of a group of plants with a common trait.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
13 July 2000 (13.07.2000) PCT  
(10) International Publication Number  
WO 00/40695 A3

- (51) International Patent Classification: C12N 15/29, C07K 14/415  
(21) International Application Number: PCT/US00/00466  
(22) International Filing Date: 7 January 2000 (07.01.2000)  
(25) Filing Language: English  
(26) Publication Language: English  
(30) Priority Data: 8 January 1999 (08.01.1999) US 60/115,293  
(71) Applicant (for all designated States except US): CERES, INC. (US); 3007 Mulibu Canyon Road, Malibu, CA 90265 (US).
- (72) Inventors: and  
(75) Inventors/Applicants (for US only): ALEXANDROV, Nikolai (RU/US); 1404 Oak Trail Street, Thousand Oaks, CA 91320 (US); BROVER, Vyacheslav (RU/US); 5916 N. Las Virgenes Road #590, Calabasas, CA 91320 (US); CHEN, Xiaofeng (CN/US); 12333 Wild Turkey Court, #B, Creeve Coeur, MO 63141 (US); SUBRAMANIAN, Gopalakrishnan (IN/US); 4205 Peach Slope Road, Moorpark, CA 93021 (US); TROUKHAN, Maxim. E. (RU/US); 1675 Amberwood Drive #2, South Pasadena, CA 91030 (US); ZHENG, Liansheng (CN/US); 19712 Circle Gate Drive #201, Germantown, MD 20874 (US).
- (74) Agents: STEWART, Raymond, C. et al.; Birch, Stewart, Kolasch & Birch, LLP, P.O. Box 747, Falls Church, VA 22040-0747 (US).
- Published: — with international search report  
(88) Date of publication of the international search report: 30 November 2000  
(48) Date of publication of this corrected version: 4 October 2001  
(15) Information about Correction: see PCT Gazette No. 40/2001 of 4 October 2001, Section II
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(64) Title: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES ENCODED THEREBY

(57) Abstract: The present invention provides DNA molecules that constitute portions of the genome of a plant, and polypeptides encoded thereby. The DNA molecules are useful for expressing a gene product, either as a promoter or as a structural gene or as an UTR or as a 3' termination sequence and are also useful in controlling expression of the target gene or as tools for genetic mapping or identification of a particular individual plant or for clustering of a group of plants with a common trait.

# SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES ENCODED THEREBY

## FIELD OF THE INVENTION

The present invention relates to isolated polynucleotides that encode all, or a portion of, a gene that is expressed and the corresponding polypeptide. The present invention also relates to isolated polynucleotides that encode regulatory regions of genes.

## DESCRIPTION OF THE RELATED ART

Efforts to map and sequence the genome of a number of organisms are in progress; a few complete genome sequences, for example those of *E. coli* and *Saccharomyces cerevisiae* are known (Blattner et al., Science 277:1453 (1997); Goffeau et al., Science 274:546 (1996)). The complete genome of a multicellular organism, *C. elegans*, has also been sequenced (See, the *C. elegans* Sequencing Consortium, Science 282:2012 (1998)). To date, no complete genome of a plant has been sequenced, nor has a complete cDNA complement of any plant been sequenced.

## SUMMARY OF THE INVENTION

The present invention comprises polynucleotides, such as complete cDNA sequences and/or sequences of genomic DNA encompassing complete genes, portions of genes, and/or intergenic regions, hereinafter collectively referred to as "Sequence-Determined DNA Fragments" (SDFs), from plants, particularly corn and *Arabidopsis thaliana* and polypeptides derived therefrom. In some instances, the SDFs span the entirety of a protein-coding segment. In some instances, the entirety of an mRNA is represented. Other objects of the invention are the control sequences, such as but not limited to promoters, that are also represented by SDFs of the invention. Complements of any sequence of the invention are also considered part of the invention.

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/00466

## Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 172(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 10, 11, 12, 22, 23  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
1-9, 13-21, 25-28

Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (continuation of first sheet(1)) (July 1998)\*

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/00466

## BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s) 1-9, 13-21, and 25-28, drawn to nucleic acids and polypeptides.

Group II, claim(s) 24, drawn to antibodies specific to the polypeptides of Group I.

The inventions listed as Groups I-II do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: PCT Rule 13.1 and Annex B do not provide for unity of invention between two different products that share a special technical feature.

In addition, each Group detailed above reads on distinct Groups drawn to multiple SEQ ID NOS. The sequences are distinct because they are unrelated sequences, and a further lack of unity is applied to each Group. The lack of unity is partially waived and the Applicants must further elect 10 SEQ ID NOS for examination in the elected Group detailed above.

Form PCT/ISA/210 (extra sheet) (July 1998)\*

2

Other objects of the invention are polynucleotides comprising exon sequences, polynucleotides comprising intron sequences, polynucleotides comprising introns together with exons, intron/exon junction sequences, 5' untranslated sequences, and 3' untranslated sequences of the SDFs of the present invention. Polynucleotides representing the joiner of any exons described herein, in any arrangement, for example, to produce a sequence encoding any desirable amino acid sequence are within the scope of the invention.

The present invention also resides in probes useful for isolating and identifying nucleic acids that hybridize to an SDF of the invention. The probes are typically of a length of 12 to 2000 nucleotides long; more typically, 15 to 200 nucleotides long; even more typically, 18 to 100 nucleotides long.

Yet another object of the invention is a method of isolating and/or identifying nucleic acids using the following steps:

(a) contacting a probe of the instant invention with a polynucleotide sample under conditions that permit hybridization and formation of a polynucleotide duplex; and

(b) detecting and/or isolating the duplex of step (a).

The conditions for hybridization can be from low to moderate to high stringency conditions. The sample can include a polynucleotide having a sequence unique in a plant genome. Probes and methods of the invention are useful, for example, without limitation, for mapping of genetic traits and/or for positional cloning of a desired portion of genomic DNA.

Probes and methods of the invention can also be used for detecting alternatively spliced messages within a species. Probes and methods of the invention can further be used to detect or isolate related genes in other plant species using genomic DNA (gDNA) and/or cDNA libraries. In some instances, especially when longer probes and low to moderate stringency

3

hybridization conditions are used, the probe will hybridize to a plurality of cDNA and/or gDNA sequences of a plant. This approach is useful for isolating representatives of gene families which are identifiable by possession of a common functional domain in the gene product or which have common cis-acting regulatory sequences. This approach is also useful for identifying orthologous genes from other organisms, which can be more or less related to corn, *Arabidopsis*, or another plant.

The present invention also resides in constructs for modulating the expression of the genes comprised of all or a portion of an SDF. The constructs comprise all or a portion of the expressed SDF, or of a complementary sequence. Examples of constructs include ribozymes comprising RNA encoded by an SDF or by a sequence complementary thereto, antisense constructs, constructs comprising coding regions or parts thereof, constructs comprising promoters, introns, untranslated regions, etc. When inserted into a host cell the construct is, preferably, functionally integrated with or operatively linked to a heterologous polynucleotide. For instance, a coding region from an SDF might be operably linked to a promoter that is functional in a plant.

The present invention also resides in host cells, including bacterial or yeast cells or plant cells, and transgenic plants that harbor constructs such as described above. Another aspect of the invention relates to methods for modulating expression of specific genes in transgenic plants by expression of the structural gene component of the constructs, by regulation of expression of one or more endogenous genes in a transgenic plant or by suppression of expression of the polynucleotides of the invention in a transgenic plant. Methods of modulation of gene expression include without limitation (1) inserting into a host cell additional copies of a polynucleotide comprising a coding sequence; (2) modulating an endogenous promoter in a host cell; (3) inserting antisense

4

or ribozyme constructs into a host cell and (4) inserting into a host cell a polynucleotide comprising a sequence encoding a mutant, fragment, or fusion of the native polypeptides of the instant invention.

5

#### BRIEF DESCRIPTION OF THE TABLES

The sequences of exemplary SDFs and polypeptides encoded thereby of the instant invention are listed in the Sequence Listing; annotation relevant to the sequences shown in the Sequence Listing is presented in TABLE 1. Each sequence corresponds to a Maximum Length cDNA Polynucleotide Sequence. Each Maximum Length cDNA Polynucleotide Sequence (Ac) corresponds to at least one sequence in the Sequence Listing. TABLE 1 corresponds with the Sequence Listing.

TABLE 1 is a Reference Table that correlates each of the sequences and SEQ ID NOS in the Sequence Listing with a corresponding Maximum Length cDNA Sequence (Ac), Ceres (Applicant) sequence identifier and other information about the individual sequence. The Sequence Listing contains the sequence of each nucleic acid and amino acid sequence.

In TABLE 1, each section begins by identifying the Maximum Length cDNA Polynucleotide Sequence, indicating a "Clone ID" that is a number used for identification purposes by the applicant and in some instances a "Public Genomic DNA" sequence, indicated by a "gi No". In those instances where a public sequence is recited, there follows information about gene annotations such as predicted exons. In this portion, after the description of the gene's functional unit, the starting and ending nucleotide number of the public sequence and the computer program used to generate the result are listed. "INIT" denotes an initial exon. "INTR" denotes an internal exon. "TERM" denotes a terminal exon. For those sequences where the computer program is identified as "OCKHAM", "INIT" denotes the first detected exon and "TERM"

5

denotes the last detected exon. "SINGLE" denotes the entire mRNA coding portion is in one exon. Other notations are considered intuitive to the skilled artisan.

In the next section (Ac), the cDNA MLS is identified by its SEQ ID NO ("Pat. Appln. SEQ ID NO") and the Ceres sequence identifier ("Ceres seq\_id"), which is also merely an identifier useful for the applicant. The designation of "Alternative transcription start sites" can include both positive and negative numbers. Positive numbers refer to the referenced SEQ ID NO directly. The positions indicated by negative numbers, if any, refer to positions in the public genomic sequences. In instances where there is a "Public Genomic DNA" referenced, the relevant genomic sequence can be found by direct reference to the nucleotide sequence indicated by the "gi" number shown for the Public Genomic DNA. In instances where there is no Public Genomic DNA referenced, the relevant nucleotide sequence for alignment is the nucleotide sequence associated with the amino acid sequence designated by a "gi" number in the section (Dp). In these instances, the nucleotide sequence is found in GENBANK by clicking on the link in the National Center for Biotechnology Information Entrez database. The numbering is relative to position 1 as determined by aligning the first residue of the MLS cDNA sequence (SEQ ID NO \*) with the genomic sequence corresponding to the relevant "gi" number.

Subsection (B) lists SEQ ID NOS and Ceres seq\_ids for polypeptide sequences encoded by the cDNA sequence and the location of the start codon within the cDNA sequence that codes for the polypeptide. Subsection (B) also describes additional features within the polypeptide sequence.

Subsection (C) provides information regarding identified domains (where present) within the polypeptide and (where present) a name for the polypeptide. Subsection (Dp) provides (where present) information concerning amino acid sequences

that are found to be related and have some sequence identity to the polypeptide sequences of the Sequence Listing. Those "related" sequences are identified by a "gi" number and are amino acid sequences in the publicly accessible BLAST databases on the NCBI FTP web site (accessible at [ncbi.nlm.gov/blast](http://ncbi.nlm.gov/blast)). The database at the NCBI FTP site utilizes the "gi" identifiers to assign by NCBI a unique identifier for each sequence in the databases, thereby providing a non-redundant database for sequences from various databases, including GenBank, EMBL, DDBJ (DNA Database of Japan) and PDB (Brookhaven Protein Data Bank). Subsection (Ba), when present, describes a sequence as being considered plant-specific (i.e. a gene found only in a plant) or describes a biochemical activity for the protein encoded by the exemplary SDF. Subsection (Dn) provides polynucleotide sequences (where present) related to the Maximum length cDNA sequence.

#### DETAILED DESCRIPTION OF THE INVENTION

The invention relates to (I) polynucleotides and methods of use thereof, such as

- IA. Probes, Primers and Substrates;
- IB. Methods of Detection and Isolation;
  - B.1. Hybridization;
  - B.2. Methods of Mapping;
  - B.3. Southern Blotting;
  - B.4. Isolating cDNA from Related Organisms;
  - B.5. Isolating and/or Identifying Orthologous Genes
- IC. Methods of Inhibiting Gene Expression
  - C.1. Antisense
  - C.2. Ribozyme Constructs;
  - C.3. Co-Suppression;
  - C.4. Other Methods to Inhibit Gene Expression
- ID. Methods of Functional Analysis;

- IE. Promoter Sequences and Their Use;
- IF. UTR's and/or Intron Sequences and Their Use; and
- IG. Coding Sequences and Their Use.

The specification also discloses (II) polypeptides including, without limitation, native proteins, mutants, fragments, and fusions. Antibodies to said polypeptides are also disclosed.

The specification also discloses (III) methods of modulating polypeptide production or activity. Examples of such methods include (i) suppressed, (ii) enhanced, and (iii) directed expression.

The specification also discloses (IV) gene constructs and construction of expression vectors, including (IVA) coding sequences and (IVB) promoters, and (IVC) Signal Peptides, (V) transformation procedures to illustrate the invention by way of examples.

#### I. Polynucleotides

A number of the nucleotide sequences disclosed in the Sequence Listing herein as representative of the SDFs of the invention can be obtained by sequencing genomic DNA (gDNA) and/or cDNA from corn plants grown from HYBRID SEED # 35A19, purchased from Pioneer Hi-Bred International, Inc., Supply Management, P.O. Box 256, Johnston, Iowa 50131-0256.

Exemplified SDFs of the invention represent portions of the genome of corn or *Arabidopsis* and/or represent mRNA expressed from that genome. The isolated nucleic acid of the invention also encompasses corresponding portions of the genome and/or cDNA complement of other organisms as described in detail below.

Starting material for cDNA synthesis for the exemplary corn cDNA clones having sequences presented in the Sequence Listing was poly(A)-containing polysomal mRNAs from



inflorescences and root tissues of corn plants grown from HYBRID SEED # 35A19. Male inflorescences and female (pre-and post-fertilization) inflorescences were isolated at various stages of development. Selection for poly(A) containing polysomal RNA was done using oligo d(T) cellulose columns, as described by Cox and Goldberg, "Plant Molecular Biology: A Practical Approach", pp. 1-35, Shaw ed., c. 1988 by IRL, Oxford.

Tissues were or each organ was individually pulverized and frozen in liquid nitrogen. Next, the samples were homogenized in the presence of detergents and then centrifuged. The debris and nuclei were removed from the sample and more detergents were added to the sample. The sample was centrifuged and the debris was removed. Then the sample was applied to a 2M sucrose cushion to isolate polysomes. The RNA was isolated by treatment with detergents and proteinase K followed by ethanol precipitation and centrifugation. The polysomal RNA from the different tissues was pooled according to the following mass ratios: 15/15/1 for male inflorescences, female inflorescences and root, respectively. The pooled material was then used for cDNA synthesis by the methods described below.

The quality and the integrity of the polyA+ RNAs were evaluated.

A number of the nucleotide sequences disclosed in the Sequence Listing herein as representative of the SDFs of the invention can also be obtained by sequencing genomic DNA from *Arabidopsis thaliana*, Wassilewskija ecotype or by sequencing cDNA obtained from mRNA from such plants as described below.

This is a true breeding strain. Seeds of the plant are available from the Arabidopsis Biological Resource Center at the Ohio State University, under the accession number CS2360. Seeds of this plant were deposited under the terms and conditions of the Budapest Treaty at the American Type Culture

Collection, Manassas, VA on August 31, 1999, and were assigned ATCC No. PTA-595.

Starting material for cDNA synthesis for the exemplary *Arabidopsis* cDNA clones having sequences presented in the Sequence Listing was polysomal RNA isolated from the top-most inflorescence tissues and roots of *Arabidopsis thaliana* Landsberg erecta (L. ex.) also obtained from the Arabidopsis Biological Resource Center. Nine parts inflorescence to every part root was used, as measured by mass. Tissue was pulverized and exposed to liquid nitrogen. Next, the sample was homogenized in the presence of detergents and then centrifuged. The debris and nuclei were removed from the sample and more detergents were added to the sample. The sample was centrifuged and the debris was removed and the sample was applied to a 2M sucrose cushion to isolate polysomal RNA. Cox et al., "Plant Molecular Biology: A Practical Approach", pp. 1-35, Shaw ed., c. 1988 by IRL, Oxford. The polysomal RNA was used for cDNA synthesis by the methods described below. Polysomal mRNA was then isolated as described above for corn cDNA. The quality of the RNA was assessed electrophoretically.

Following preparation of the mRNAs from various tissues as described above, selection of mRNA with intact 5' ends and specific attachment of an oligonucleotide tag to the 5' end of such mRNA was performed using either a chemical or enzymatic approach. Both techniques take advantage of the presence of the "cap" structure, which characterizes the 5' end of most intact mRNAs and which comprises a guanosine generally methylated once, at the 7 position.

The chemical modification approach involves the optional elimination of the 2', 3'-cis diol of the 3' terminal ribose, the oxidation of the 2', 3'-cis diol of the ribose linked to the cap of the 5' ends of the mRNAs into a dialdehyde, and the coupling of the such obtained dialdehyde to a derivatized

10

oligonucleotide tag. Further detail regarding the chemical approaches for obtaining mRNAs having intact 5' ends are disclosed in International Application No. WO96/34981 published November 7, 1996.

5 The enzymatic approach for ligating the oligonucleotide tag to the intact 5' ends of mRNAs involves the removal of the phosphate groups present on the 5' ends of uncapped incomplete mRNAs, the subsequent decapping of mRNAs having intact 5' ends and the ligation of the phosphate present at the 5' end of the decapped mRNA to an oligonucleotide tag. Further detail regarding the enzymatic approaches for obtaining mRNAs having intact 5' ends are disclosed in Dumas Milne Edwards J.B. (Doctoral Thesis of Paris VI University, Le clonage des ADNC complets: difficultes et perspectives nouvelles. Apports pour 15 l'etude de la regulation de l'expression de la tryptophane hydroxylase de rat, 20 Dec. 1993), EPO 625572 and Kato et al., Gene 150:243-250 (1994).

In both the chemical and the enzymatic approach, the oligonucleotide tag has a restriction enzyme site (e.g. an EcoRI site) therein to facilitate later cloning procedures. Following attachment of the oligonucleotide tag to the mRNA, the integrity of the mRNA is examined by performing a Northern blot using a probe complementary to the oligonucleotide tag.

11

For the mRNAs joined to oligonucleotide tags using either the chemical or the enzymatic method, first strand cDNA synthesis is performed using an oligo-dT primer with reverse transcriptase. This oligo-dT primer can contain an internal tag of at least 4 nucleotides, which can be different from one mRNA preparation to another. Methylated dCTP is used for cDNA first strand synthesis to protect the internal EcoRI sites from digestion during subsequent steps. The first strand cDNA is precipitated using isopropanol after removal of RNA by alkaline hydrolysis to eliminate residual primers.

Second strand cDNA synthesis is conducted using a DNA polymerase, such as Klenow fragment and a primer corresponding to the 5' end of the ligated oligonucleotide. The primer is typically 20-25 bases in length. Methylated dCTP is used for second strand synthesis in order to protect internal EcoRI sites in the cDNA from digestion during the cloning process.

Following second strand synthesis, the full-length cDNAs are cloned into a phagemid vector, such as pBlueScript™ (Stratagene). The ends of the full-length cDNAs are blunted with T4 DNA polymerase (Biolabs) and the cDNA is digested with EcoRI. Since methylated dCTP is used during cDNA synthesis, the EcoRI site present in the tag is the only hemi-methylated site; hence the only site susceptible to EcoRI digestion. In some instances, to facilitate subcloning, an Hind III adapter is added to the 3' end of full-length cDNAs.

The full-length cDNAs are then size fractionated using either exclusion chromatography (ACh, Biosepra) or electrophoretic separation which yields 3 to 6 different fractions. The full-length cDNAs are then directionally cloned either into pBlueScript™ using either the EcoRI and SmaI restriction sites or, when the Hind III adapter is present in the full-length cDNAs, the EcoRI and Hind III restriction sites. The ligation mixture is transformed, preferably by

12

electroporation, into bacteria, which are then propagated under appropriate antibiotic selection.

Clones containing the oligonucleotide tag attached to full-length cDNAs are selected as follows.

5 The plasmid cDNA libraries made as described above are purified (e.g. by a column available from Qiagen). A positive selection of the tagged clones is performed as follows. Briefly, in this selection procedure, the plasmid DNA is converted to single stranded DNA using phage F1 gene II endonuclease in combination with an exonuclease (Chang et al., Gene 127:95 (1993)) such as exonuclease III or T7 gene 6 exonuclease. The resulting single stranded DNA is then purified using paramagnetic beads as described by Fry et al., *Biotechniques* 13: 124 (1992). Here the single stranded DNA is 15 hybridized with a biotinylated oligonucleotide having a sequence corresponding to the 3' end of the oligonucleotide tag. Preferably, the primer has a length of 20-25 bases. Clones including a sequence complementary to the biotinylated oligonucleotide are selected by incubation with streptavidin 20 coated magnetic beads followed by magnetic capture. After capture of the positive clones, the plasmid DNA is released from the magnetic beads and converted into double stranded DNA using a DNA polymerase such as ThermoSequenase™ (obtained from Amersham Pharmacia Biotech). Alternatively, protocols such as 25 the Gene Trapper™ kit (Gibco BRL) can be used. The double stranded DNA is then transformed, preferably by electroporation, into bacteria. The percentage of positive clones having the 5' tag oligonucleotide is typically estimated to be between 90 and 98% from dot blot analysis.

Following transformation, the libraries are ordered in microtiter plates and sequenced. The Arabidopsis library was deposited at the American Type Culture Collection on January 7, 2000 as "E-coli liba 010600" under the accession number \_

13

Other methods for cloning full-length cDNA are described, for example, by Seki et al., *Plant Journal* 15:707-720 (1998) "High-efficiency cloning of Arabidopsis full-length cDNA by biotinylated Cap trapper"; Maruyama et al., Gene 138:171 (1994) "Oligo-capping a simple method to replace 5 the cap structure of eukaryotic mRNAs with oligoribonucleotides"; and WO 96/34981.

It is contemplated that the nucleotide sequences presented herein may contain some small percentage of errors. These errors may arise in the normal course of determination of 10 nucleotide sequences. Sequence errors can be corrected by obtaining seeds deposited under the accession numbers cited above, propagating them, isolating genomic DNA or appropriate mRNA from the resulting plants or seeds thereof, amplifying the relevant portion of the genomic DNA or mRNA using primers 15 having a sequence that flanks the erroneous sequence, and sequencing the amplification product.

#### I.A. Probes, Primers and Substrates

SDFs of the invention can be applied to substrates for 20 use in array applications such as, but not limited to, assays of global gene expression, for example under varying conditions of development, growth conditions. The arrays can also be used in diagnostic or forensic methods.

Probes and primers of the instant invention will 25 hybridize to a polynucleotide comprising a sequence in the Sequence Listing. Though many different nucleotide sequences can encode an amino acid sequence, in some instances, the sequences of the Sequence Listing are preferred for encoding polypeptides of the invention. However, the sequence of the 30 probes and/or primers of the instant invention need not be identical to those in the Sequence Listing or the complements thereof. For example, some variation in probe or primer sequence and/or length can allow additional family members to

14

be detected, as well as orthologous genes and more taxonomically distant related sequences. Similarly probes and/or primers of the invention can include additional nucleotides that serve as a label for detecting the formed duplex or for subsequent cloning purposes.

Probe length will vary depending on the application. For use as PCR primers, probes should be 12-40 nucleotides, preferably 18-30 nucleotides long. For use in mapping, probes should be 50 to 500 nucleotides, preferably 100-250 nucleotides long. For Southern hybridizations, probes as long as several kilobases can be used as explained below.

The probes and/or primers can be produced by synthetic procedures such as the triester method of Matteucci et al. *J. Am. Chem. Soc.* 103:3185 (1981); or according to Urdea et al. *Proc. Natl. Acad. Sci.* 80:7461 (1981) or using commercially available automated oligonucleotide synthesizers.

### I.B. Methods of Detection and Isolation

#### B.1. Hybridization

Probes and/or primers can be used for detection and/or isolation of polynucleotide sequences. Such polynucleotides are included in the subject matter of the invention. Depending on the stringency of the conditions under which these probes and/or primers are used, polynucleotides exhibiting a wide range of similarity to those in the Sequence Listing can be detected or isolated.

"Stringency" is a function of probe length, probe composition (G + C content), and salt concentration, organic solvent concentration, and temperature of hybridization or wash conditions. Stringency is typically compared by the parameter " $T_m$ ", which is the temperature at which 50% of the complementary molecules in the hybridization are hybridized, in terms of a temperature differential from  $T_m$ . High stringency

15

conditions are those providing a condition of  $T_m$  - 5°C to  $T_m$  - 10°C. Medium stringency conditions are those providing  $T_m$  - 20°C to  $T_m$  - 29°C. Low stringency conditions are those providing a condition of  $T_m$  - 40°C to  $T_m$  - 48°C. The relationship of hybridization conditions to  $T_m$  (in °C) is expressed in the mathematical equation

$$T_m = 81.5 - 16.6(\log_{10}[\text{Na}^+]) + 0.41(\%G+C) - (600/N) \quad (1)$$

where N is the length of the probe. This equation works well for probes 14 to 70 nucleotides in length that are identical to the target sequence. The equation below for  $T_m$  of DNA-DNA hybrids is useful for probes in the range of 50 to greater than 500 nucleotides, and for conditions that include an organic solvent (formamide).

$$T_m = 81.5 + 16.6 \log \{ [\text{Na}^+] / (1 + 0.7[\text{Na}^+]) \} + 0.41(\%G+C) - 500/L \cdot 0.63(\% \text{formamide}) \quad (2)$$

where L is the length of the probe in the hybrid. (P. Tijessen, "Hybridization with Nucleic Acid Probes" in Laboratory Techniques in Biochemistry and Molecular Biology, P.C. van der Vliet, ed., c. 1993 by Elsevier, Amsterdam.) The  $T_m$  of equation (2) is affected by the nature of the hybrid; for DNA-RNA hybrids  $T_m$  is 10-15°C higher than calculated, for RNA-RNA hybrids  $T_m$  is 20-25°C higher. Because the  $T_m$  decreases about 1°C for each 1% decrease in homology when a long probe is used (Bonner et al., *J. Mol. Biol.* 81:123 (1973)), stringency conditions can be adjusted to favor detection of identical genes or related family members.

Equation (2) is derived assuming equilibrium and therefore, hybridizations according to the present invention are most preferably performed under conditions of probe excess and for sufficient time to achieve equilibrium. The time required to reach equilibrium can be shortened by inclusion of a "hybridization accelerator" such as dextran

16

sulfate or another high volume polymer in the hybridization buffer.

When the practitioner wishes to examine the result of membrane hybridizations under a variety of stringencies, an efficient way to do so is to perform the hybridization under a low stringency condition, then to wash the hybridization membrane under increasingly stringent conditions. With respect to wash steps, preferred stringencies lie within the ranges stated above; high stringency is 5-8°C below  $T_m$ , medium stringency is 26-29°C below  $T_m$  and low stringency is 45-48°C below  $T_m$ .

A number of methods known to those skilled in the art can be used with the probes and/or primers of the invention to isolate and detect polynucleotides, including, without limitation: Southern, Northern, Branched DNA hybridization assays, polymerase chain reaction, and variations thereof.

When using SDFs to identify orthologous genes in other species, the practitioner will preferably adjust the amount of target DNA of each species so that, as nearly as is practical, the same number of genome equivalents are present for each species examined. This prevents faint signals from species having large genomes, and thus small numbers of genome equivalents per mass of DNA, from erroneously being interpreted as absence of the corresponding gene in the genome.

A good general discussion of the factors for determining hybridization conditions is provided by Sambrook et al. ("Molecular Cloning, a Laboratory Manual, 2nd ed., c. 1989 by Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; see esp., chapters 11 and 12). Additional considerations and details of the physical chemistry of hybridization are provided by G.H. Keller and M.M. Manak "DNA Probes", 2<sup>nd</sup> Ed. pp. 1-25, c. 1993 by Stockton Press, New York, NY.

17

Hybridization of one nucleic acid to another constitutes a physical property that defines the subject SDF of the invention. Also, such hybridization imposes structural limitations on the pair. For example, for a probe molecule, given that the sequence of the probe nucleic acid is known and fixed, equation (2) indicates that the combined variation in GC content of the target DNA and mismatch between the probe and the hybridizing DNA is determined for any given hybridization buffer composition and  $T_m$ .

The probes and/or primers of the instant invention can be used to detect or isolate nucleotides that are "identical" to the probes or primers. Two nucleic acid sequences or polypeptides are said to be "identical" if the sequence of nucleotides or amino acid residues, respectively, in the two sequences is the same when aligned for maximum correspondence as described below. The term "complementary to" is used herein to mean that the sequence can form a Watson-Crick base pair with a reference polynucleotide sequence. Complementary sequences can include nucleotides, such as inosine, that neither disrupt Watson-Crick base pairing nor contribute to the pairing.

Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman *Add. APL. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman and Wunsch *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson and Lipman *Proc. Natl. Acad. Sci. (USA)* 85: 2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, PASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection. Given that two sequences have been identified for comparison, GAP and BESTFIT are preferably employed to determine their optimal alignment. Typically, the

18

default values of 5.00 for gap weight and 0.30 for gap weight length are used.

The probes and/or primers of the invention can also be used to detect and/or isolate polynucleotides exhibiting at least 80% sequence identity with the sequences of the Sequence Listing or fragments thereof.

"Percentage of sequence identity" is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (e.g., gaps or overhangs) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

"Percentage of sequence identity" can be determined by the algorithms described above.

The term "substantial identity" between polynucleotide or polypeptide sequences refers to polynucleotide or polypeptide comprising a sequence that has at least 80% sequence identity, preferably at least 85%, more preferably at least 90% and most preferably at least 95%, even more preferably, at least 96%, 97%, 98% or 99% sequence identity compared to a reference sequence using the programs.

Isolated polynucleotides within the scope of the invention also include allelic variants of the specific sequences presented in the Sequence Listing. An "allelic variant" is a sequence that is a variant from that of the SDF, but represents the same chromosomal locus in the organism. Allelic variants can arise by normal genetic

19

variation in a population. Allelic variants can also be produced by genetic engineering methods. An allelic variant can be one that is found in a naturally occurring plant, including a cultivar or ecotype. A silent allele can give rise to phenotypic and expression profiles. An allelic variant may or may not give rise to a phenotypic change, and may or may not be expressed. An expressed allele can result in a detectable change in the phenotype of the trait represented by the locus. Allelic variations can occur in any portion of the gene sequence, including regulatory regions as well as structural regions.

With respect to nucleotide sequences, degeneracy of the genetic code provides the possibility to substitute at least one base of the base sequence of a gene with a different base without causing the amino acid sequence of the polypeptide produced from the gene to be changed. Hence, the DNA of the present invention may also have any base sequence that has been changed from a sequence in the Sequence Listing by substitution in accordance with degeneracy of genetic code.

References describing codon usage include: Carrels et al., *J. Mol. Evol.* 46: 45 (1998) and Fennoy et al., *Nucl. Acids Res.* 21(23): 5294 (1993).

#### A. B.2. Mapping

The isolated SDF DNA of the invention can be used to create various types of genetic and physical maps of the genome of corn, Arabidopsis or other plants. Some SDFs may be absolutely associated with particular phenotypic traits, allowing construction of gross genetic maps. While not all SDFs will immediately be associated with a phenotype, all SDFs can be used as probes for identifying polymorphisms associated with phenotypes of interest. Briefly, total DNA is isolated from individuals and is subsequently cleaved with one or more restriction enzymes, separated according to mass,

20

transferred to a solid support, hybridized with SDF DNA and the pattern of fragments compared. Polymorphisms associated with a particular SDF are visualized as differences in the size of fragments produced between individual DNA samples after digestion with a particular restriction enzyme and hybridization with the SDF. After identification of polymorphic SDF sequences, linkage studies can be conducted. Recombinants produced are analyzed using the same restriction enzyme/hybridization procedure. After identification of many polymorphisms using SDF sequences, linkage studies can be conducted by using the individuals showing polymorphisms as parents in crossing programs. F2 progeny recombinants or recombinant inbreds, for example, are then analyzed using the same restriction enzyme/hybridization procedure. The order of DNA polymorphisms along the chromosomes can be inferred based on the frequency with which they are inherited together versus independently. The closer two polymorphisms are together in a chromosome the higher the probability that they are inherited together. Integration of the relative positions of all the polymorphisms and associated marker SDFs produces a genetic map of the species, where the distances between markers reflect the recombination frequencies in that chromosome segment.

The use of recombinant inbred lines for such genetic mapping is described for *Arabidopsis* by Alonso-Blanco et al. (*Methods in Molecular Biology*, vol. 82, "Arabidopsis Protocols", pp. 137-146, J.M. Martinez-Zapater and J. Salinas, eds., c. 1998 by Humana Press, Totowa, NJ) and for corn by Burr ("Mapping Genes with Recombinant Inbreds", pp. 249-254. In Freeling, M. and V. Walbot (Ed.), *The Maize Handbook*, c. 1994 by Springer-Verlag New York, Inc.: New York, NY, USA; Berlin Germany; Burr et al. *Genetics* (1998) 118: 519; Gardiner, J. et al., (1993) *Genetics* 134: 917). However, this procedure is not limited to plants and can be

21

used for other organisms (such as yeast) or for individual cells.

The SDFs of the present invention can also be used for simple sequence repeat (SSR) mapping. Rice SSR mapping is described by Morgante et al. (*The Plant Journal* (1993) 3: 165), Panaud et al. (*Genome* (1995) 38: 1170); Senior et al. (*Crop Science* (1996) 36: 1676), Taramino et al. (*Genome* (1996) 39: 277) and Ahn et al. (*Molecular and General Genetics* (1993) 241: 483-90). SSR mapping can be achieved using various methods. In one instance, polymorphisms are identified when sequence specific probes flanking an SSR contained within an SDF are made and used in polymerase chain reaction (PCR) assays with template DNA from two or more individuals of interest. Here, a change in the number of tandem repeats between the SSR-flanking sequence produces differently sized fragments (U.S. Patent 5,766,847). Alternatively, polymorphisms can be identified by using the PCR fragment produced from the SSR-flanking sequence specific primer reaction as a probe against Southern blots representing different individuals (U.H. Refseth et al., (1997) *Electrophoresis* 18: 1519).

Genetic and physical maps of crop species have many uses. For example, these maps can be used to devise positional cloning strategies for isolating novel genes from the mapped crop species. In addition, because the genomes of closely related species are largely syntenic (that is, they display the same ordering of genes within the genome), these maps can be used to isolate novel alleles from wild relatives of crop species by positional cloning strategies.

The various types of maps discussed above can be used with the SDFs of the invention to identify Quantitative Trait Loci (QTLs). Many important crop traits, such as the solids content of tomatoes, are quantitative traits and result from the combined interactions of several genes. These genes

reside at different loci in the genome, oftentimes on different chromosomes, and generally exhibit multiple alleles at each locus. The SDFs of the invention can be used to identify QTLs and isolate specific alleles as described by de Vicente and Tanksley (*Genetics* 134:585 (1993)). In addition to isolating QTL alleles present in crop species, the SDFs of the invention can also be used to isolate alleles from the corresponding QTL of wild relatives. Transgenic plants having various combinations of QTL alleles can then be created and the effects of the combinations measured. Once an ideal allele combination has been identified, crop improvement can be accomplished either through biotechnological means or by directed conventional breeding programs (for review see Tanksley and McCouch, *Science* 277:1063 (1997)).

In another embodiment the SDFs can be used to help create physical maps of the genome of corn, *Arabidopsis* and related species. Where SDFs have been ordered on a genetic map, as described above, then SDFs can be used as probes to discover which clones in large libraries of plant DNA fragments in YACs, BACs, etc. contain the same SDF or similar sequences, thereby facilitating the assignment of the large DNA fragments to chromosomal positions. Subsequently, the large BACs, YACs, etc. can be ordered unambiguously by more detailed studies of their sequence composition (e.g. Maria et al. (1997) *Genomic Research* 7:1072-1084) and by using their end or other sequences to find the identical sequences in other cloned DNA fragments. The overlapping of DNA sequences in this way allows large contigs of plant sequences to be built, that, when sufficiently extended, provide a complete physical map of a chromosome. Sometimes the SDFs themselves will provide the means of joining cloned sequences into a contig.

The patent publication WO95/35505 and U.S. Patents 5,445,943 and 5,410,270 describe scanning multiple alleles of a plurality of loci using hybridization to arrays of oligonucleotides. These techniques are useful for each of the types of mapping discussed above.

Following the procedures described above and using a plurality of the SDFs of the present invention, any individual can be genotyped. These individual genotypes can be used for the identification of particular cultivars, varieties, lines, ecotypes and genetically modified plants or can serve as tools for subsequent genetic studies involving multiple phenotypic traits.

#### B. 3 Southern Blot Hybridization

The sequences from the Sequence Listing can be used as probes for various hybridization techniques. These techniques are useful for detecting target polynucleotides in a sample or for determining whether transgenic plants, seeds or host cells harbor a gene or sequence of interest and thus might be expected to exhibit a particular trait or phenotype.

In addition, the hybridization of the SDFs of the invention to nucleic acids obtained from other organisms can be used to identify orthologous genes from other species and/or additional members of gene families either in the same or different species. In regard to identifying genes in other species, a Southern blot of genomic DNA provides description of isolated DNA fragments that comprise the orthologous genes or additional members of the gene families. That is, given such data, one of ordinary skill in the art could distinguish the isolated DNA fragments by their size together with the restriction sites at each end and by the property of hybridizing with the SDF probe under the stated conditions.



24

In addition, the SDFs from the invention can be used to isolate additional members of gene families from the same species and/or orthologous genes from different species. This is accomplished by hybridizing an SDF to a Southern blot containing the appropriate genomic DNA or cDNA. Given the resulting hybridization data, one of ordinary skill in the art could distinguish and isolate the correct DNA fragments by size, restriction sites and stated hybridization conditions from a gel or from a library.

10

Identification and isolation of orthologous genes from closely related species and alleles within a species is particularly desirable because of their potential for crop improvement. Many important crop traits, such as the solid content of tomatoes, result from the combined interactions of the products of several genes residing at different loci in the genome. Generally, alleles at each of these loci can make quantitative differences to the trait. By identifying and isolating numerous alleles for each locus from within or different species, transgenic plants having various combinations of alleles can be created and the effects of the combinations measured. Once a more favorable ideal allele combination has been identified, crop improvement can be accomplished either through biotechnological means or by directed conventional breeding programs (Tanksley et al. *Science* 277:1063(1997)).

25

The results from hybridizations of the SDFs of the invention to Southern blots containing DNA from another species can also be used to generate restriction fragment maps for the corresponding genomic regions. These maps provide map ~~provides~~ additional information about the relative positions of restriction sites within fragments, further distinguishing mapped DNA from the remainder of the genome.

30

25

Physical maps can be made by digesting genomic DNA with different combinations of restriction enzymes.

Probes for Southern blotting to distinguish individual restriction fragments can range in size from 15 to 20 nucleotides to several thousand nucleotides. More preferably, the probe is 100 to 1000 nucleotides long for identifying members of a gene family when it is found that repetitive sequences would complicate the hybridization. For identifying an entire corresponding gene in another species, the probe is more preferably the length of the gene, typically 2000 to 10,000 nucleotides, but probes 50-1,000 nucleotides long might be used. Some genes, however, might require probes up to 15,000 nucleotides long or overlapping probes constituting the full-length sequence to span their lengths.

15

Also, while it is preferred that the probe be homogeneous with respect to its sequence, that is not necessary. For example, as described below, a probe representing members of a gene family having diverse sequences can be generated using PCR to amplify genomic DNA or RNA templates using primers derived from SDFs that include sequences that define the gene family.

20

For identifying corresponding genes in another species, the probe for Southern blotting most preferably would be the genomic copy of the probe gene. This allows all elements of the gene to be identified in the other species. The next most preferable probe is a cDNA spanning the entire coding sequence, which allows all of the mRNA-coding portion of the gene to be identified; in this case it is possible that some introns in the gene might be missed. Probes for Southern blotting can easily be generated from SDFs by making primers having the sequence at the ends of the SDF and using corn or *Arabidopsis* genomic DNA as a template. In instances where the SDF includes sequence conserved among species, primers

25

30

26

including the conserved sequence can be used for PCR with genomic DNA from a species of interest to obtain a probe. Similarly, if the SDF includes a domain of interest, that portion of the SDF can be used to make primers and, with appropriate template DNA, used to make a probe to identify genes containing the domain. Alternatively, the PCR products can be resolved, for example by gel electrophoresis, and cloned and/or sequenced. In this manner, the variants of the domain among members of a gene family, both within and across species, can be examined.

#### B.4.1 Isolating DNA from Related Organisms

The SDFs of the invention can be used to isolate the corresponding DNA from other organisms. Either cDNA or genomic DNA can be isolated. For isolating genomic DNA, a lambda, cosmid, BAC or YAC, or other large insert genomic library from the plant of interest can be constructed using standard molecular biology techniques as described in detail by Sambrook et al. 1989 (Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed. Cold Spring Harbor Laboratory Press, New York) and by Ausubel et al. 1992 (Current Protocols in Molecular Biology, Greene Publishing, New York).

To screen a phage library, recombinant lambda clones are plated out on appropriate bacterial medium using an appropriate *E. coli* host strain. The resulting plaques are lifted from the plates using nylon or nitrocellulose filters. The plaque lifts are processed through denaturation, neutralization, and washing treatments following the standard protocols outlined by Ausubel et al. (1992). The plaque lifts are hybridized to either radioactively labeled or non-radioactively labeled SDF DNA at room temperature for about 16 hours, usually in the presence of 50% formamide and 5X SSC (sodium chloride and sodium citrate) buffer and blocking reagents. The plaque lifts are then washed at 42°C with 1%

27

Sodium Dodecyl Sulfate (SDS) and at a particular concentration of SSC. The SSC concentration used is dependent upon the stringency at which hybridization occurred in the initial Southern blot analysis performed. For example, if a fragment hybridized under medium stringency (e.g.,  $T_m - 20^\circ\text{C}$ ), then this condition is maintained or preferably adjusted to a less stringent condition (e.g.,  $T_m - 30^\circ\text{C}$ ) to wash the plaque lifts. Positive clones show detectable hybridization e.g., by exposure to X-ray films or chromogen formation. The positive clones are then subsequently isolated for purification using the same general protocol outlined above. Once the clone is purified, restriction analysis can be conducted to narrow the region corresponding to the gene of interest. The restriction analysis and succeeding subcloning steps can be done using procedures described by, for example Sambrook et al. (1989) cited above.

To screen a YAC library, the procedures outlined for the lambda library are essentially similar except the YAC clones are harbored in bacterial colonies. The YAC clones are plated out at reasonable density on nitrocellulose or nylon filters supported by appropriate bacterial medium in petri plates. Following the growth of the bacterial clones, the filters are processed through the denaturation, neutralization, and washing steps following the procedures of Ausubel et al. 1992. The same hybridization procedures for lambda library screening are followed.

To isolate cDNA, similar procedures using appropriately modified vectors are employed. For instance, the library can be constructed in a lambda vector appropriate for cloning cDNA such as  $\lambda$ gt11. Alternatively, the cDNA library can be made in a plasmid vector. cDNA for cloning can be prepared by any of the methods known in the art, but is preferably

28

prepared as described above. Preferably, a cDNA library will include a high proportion of full-length clones.

#### B. 5. Isolating and/or Identifying Orthologous Genes

Probes and primers of the invention can be used to identify and/or isolate polynucleotides related to those in the Sequence Listing. Related polynucleotides are those that are native to other plant organisms and exhibit either similar sequence or encode polypeptides with similar biological activity. One specific example is an orthologous gene, a gene that has a high degree of sequence similarity, often along the entire length of the coding portion of the gene, and also encodes a gene product that performs a similar function in the organism. Orthologous genes may be distinguished from homologous genes in that homologous genes share sequence similarity but often only in a portion of the sequence, which often represents a functional domain such as a tyrosine kinase activity, a DNA binding domain, or the like. The functional activities of homologous genes are not necessarily the same, but are the same for orthologous genes. The degree of identity is a function of evolutionary separation and, in closely related species, the degree of identity can be 98 to 100%. The amino acid sequence of a protein encoded by an orthologous gene can be less than 75% identical, but tends to be at least 75% or at least 80% identical, more preferably at least 90%, most preferably at least 95% identical to the amino acid sequence of the reference protein.

To find orthologous genes, the probes are hybridized to nucleic acids from a species of interest under low stringency conditions and blots are then washed under conditions of increasing stringency. It is preferable that the wash stringency be such that sequences that are 85 to 100% identical will hybridize. More preferably, sequences 90 to 100% identical will hybridize and most preferably only sequences

29

greater than 95% identical will hybridize. The low stringency condition is preferably one where sequences containing as much as 40-45% mismatches will be able to hybridize. This condition is established by  $T_m - 40^\circ\text{C}$  to  $T_m - 48^\circ\text{C}$  (see below). One of ordinary skill in the art will recognize that, due to degeneracy in the genetic code, amino acid sequences that are identical can be encoded by DNA sequences as little as 67% identical. Thus, it is preferable to make an overlapping series of shorter probes, on the order of 24 to 45 nucleotides, and individually hybridize them to the same arrayed library to avoid the problem of degeneracy introducing large numbers of mismatches.

As evolutionary divergence increases, genome sequences also tend to diverge. Thus, one of skill will recognize that searches for orthologous genes between more divergent species will require the use of lower stringency conditions compared to searches between closely related species. Also, degeneracy is more of a problem for searches in the genome of a species more distant evolutionarily from the species that is the source of the SDF probe sequences.

Therefore the method described in Bouckaert et al., U.S. Ser. No. 60/121,700 Atty. Dkt. No. 2750-117P, Client Dkt. No. 00010.001, filed February 25, 1999, hereby incorporated in its entirety by reference, can be applied to the SDFs of the present invention to isolate related genes from plant species which do not hybridize to the corn or *Arabidopsis* sequences of the Sequence Listing.

Identification of the relationship of nucleotide or amino acid sequences among plant species can be done by comparison of the subject nucleotide or amino acid sequence to the sequences of SDFs of the present application presented in the Sequence Listing.

The SDFs of the invention can also be used as probes to search for genes that are related to the SDF within a

30

species. Such related genes are typically considered to be members of a "gene family." In such a case, the sequence similarity will often be concentrated into one or a few portions of the sequence. The portions of similar sequence that define the gene family typically encode a portion of a protein or RNA that has an enzymatic or structural function. The degree of identity in the amino acid sequence of the domain that defines the gene family is preferably at least 70%, more preferably 80 to 95%, most preferably 85 to 99%.

To search for members of a gene family within a species, a "low stringency" hybridization is usually performed, but this will depend upon the size, distribution and degree of sequence divergence of domains that define the gene family. SDFs encompassing regulatory regions can be used to identify "coordinately expressed" genes by using the regulatory region portion of the SDF as a probe.

In the instances where the SDFs are identified as being expressed from genes that confer a particular phenotype, then the SDFs can also be used as probes to assay plants of different species for those phenotypes.

#### I.C. Methods to Inhibit Gene Expression

In some instances it is desirable to suppress expression of an endogenous or exogenous gene. A well-known instance is the FLAVOR-FAVOR™ tomato, in which the gene encoding ACC synthase is inactivated by an antisense approach, thus delaying softening of the fruit after ripening. See for example, U.S. Patent No. 5,859,330; U.S. Patent No. 5,723,766; Oeller, et al, Science, 254:437-439(1991); and Hamilton et al, Nature, 346:284-287 (1990). Also, timing of flowering can be controlled by suppression of the *FLOWERING LOCUS C*; high levels of this transcript are associated with late flowering, while absence of *FLC* is associated with early flowering (S.D. Michaels et al., Plant Cell 11:949 (1999)).

31

Also, the transition of apical meristem from production of leaves with associated shoots to flowering is regulated by *TERMINAL FLOWER1*, *APETALA1* and *LEAFY*. Thus, when it is desired to induce a transition from shoot production to flowering, it is desirable to suppress *TFL1* expression (S.J. Liljegren, Plant Cell 11:1007 (1999)). As another instance, it has been found that suppression of the ethylene forming enzyme results in arrested ovule development and female sterility that can be reversed by application of ethylene (D. De Martinis et al., Plant Cell 11:1061 (1999)). The ability to manipulate fertility of female plants is useful in the increasing fruit production and creating hybrids.

In an instance when it is desired to express a dominant negative mutation, it often helpful to suppress expression of the endogenous, native protein. Expression of dominant negative mutant proteins is a useful tool for research, for example when a dominant negative mutation of a receptor is used to constitutively activate or suppress a signal transduction cascade, allowing examination of the phenotype and thus the trait(s) controlled by that receptor and pathway.

#### C.1 Antisense

In the case of polynucleotides used to inhibit expression of an endogenous gene, the introduced sequence need not be perfectly identical to a sequence of the target endogenous gene. The introduced polynucleotide sequence will typically be at least substantially identical (as determined above) to the target endogenous sequence.

Some polynucleotide SDFs in the Sequence Listing represent sequences that are expressed in corn and/or *Arabidopsis*. Thus the invention includes using these sequences to generate antisense constructs to inhibit transcription and/or translation of said SDFs, typically in a plant cell.

32

To accomplish this, a polynucleotide segment from the desired gene that can hybridize to the mRNA expressed from the desired gene (the "antisense segment") is operably linked to a promoter such that the antisense strand of RNA will be transcribed when the construct is present in a host cell. A regulated promoter can be used in the construct to control transcription of the antisense segment so that transcription occurs only under desired circumstances.

The antisense segment to be introduced generally will be substantially identical to at least a portion of the endogenous gene or genes to be repressed. The sequence, however, need not be perfectly identical to inhibit expression. Further, the antisense product may hybridize to the untranslated region instead of or in addition to the coding portion of the gene. The vectors of the present invention can be designed such that the inhibitory effect applies to other proteins within a family of genes exhibiting homology or substantial homology to the target gene.

For antisense suppression, the introduced antisense segment sequence also need not be full length relative to either the primary transcription product or fully processed mRNA. Generally, higher sequence identity can be used to compensate for the use of a shorter sequence. Furthermore, the introduced sequence need not have the same intron or exon pattern, and homology of non-coding segments may be equally effective. Normally, a sequence of between about 30 or 40 nucleotides and the full length of the transcript should be used, though a sequence of at least about 100 nucleotides is preferred, a sequence of at least about 200 nucleotides is more preferred, and a sequence of at least about 500 nucleotides is especially preferred.

## C.2. Ribozymes

33

It is also contemplated that gene constructs representing ribozymes and based on the SDFs in the Sequence Listing are an object of the invention. Ribozymes can also be used to inhibit expression of genes by suppressing the translation of the mRNA into a polypeptide. It is possible to design ribozymes that specifically pair with virtually any target RNA and cleave the phosphodiester backbone at a specific location, thereby functionally inactivating the target RNA. In carrying out this cleavage, the ribozyme is not itself altered, and is thus capable of recycling and cleaving other molecules, making it a true enzyme. The inclusion of ribozyme sequences within antisense RNAs confers RNA-cleaving activity upon them, thereby increasing the activity of the constructs.

A number of classes of ribozymes have been identified. One class of ribozymes is derived from a number of small circular RNAs, which are capable of self-cleavage and replication in plants. The RNAs replicate either alone (viroid RNAs) or with a helper virus (satellite RNAs). Examples include RNAs from avocado sunblotch viroid and the satellite RNAs from tobacco ringspot virus, lucerne transient streak virus, velvet tobacco mottle virus, solanum nodiflorum mottle virus and subterranean clover mottle virus. The design and use of target RNA-specific ribozymes is described in Haselhoff et al. *Nature*, 334:585 (1988).

Like the antisense constructs above, the ribozyme sequence portion necessary for pairing need not be identical to the target nucleotides to be cleaved, nor identical to the sequences in the Sequence Listing. Generally, the sequence in the ribozyme capable of binding to the target sequence exhibits substantial sequence identity to a sequence in the Sequence Listing or the complement thereof, or to a portion of said sequence or complement. Further, the ribozyme sequence also need not be full length relative to either the primary transcription product or fully processed mRNA. The

34

ribozyme can be equally effective in inhibiting mRNA translation by cleaving either in the untranslated or coding regions. Generally, higher sequence identity can be used to compensate for the use of a shorter sequence. Furthermore, the introduced sequence need not have the same intron or exon pattern, and homology of non-coding segments may be equally effective.

### C.3. Sense Suppression

Another method of suppression is by introducing an exogenous copy of the gene to be suppressed. Introduction of expression cassettes in which a nucleic acid is configured in the sense orientation with respect to the promoter into the chromosome of a plant or by a self-replicating virus has been shown to be an effective means by which to induce degradation of mRNAs of target genes. For an example of the use of this method to modulate expression of endogenous genes see, Napoli et al., *The Plant Cell* 2:279 (1990), and U.S. Patents Nos. 5,034,323, 5,231,020, and 5,283,184. Generally, where inhibition of expression is desired, some transcription of the introduced sequence is probably necessary. The effect may occur where the introduced sequence contains no coding sequence *per se*, but comprises only intron or untranslated sequences homologous to sequences present in the primary transcript of the endogenous sequence. The introduction of only regulatory promoter sequences can also cause interference with the activity of endogenous promoters possessing the same sequence. Thus, the described SDFs can also be used to control transcription. In all of these procedures, the introduced sequence generally will be substantially identical to the endogenous sequence intended to be inactivated. The minimal identity will typically be greater than about 65%, but a higher identity might exert a more effective repression of expression of the endogenous sequences. Sequence identity of more than

35

about 80% is preferred, though about 95% to absolute identity would be most preferred. As with antisense regulation, the effect would likely apply to any other proteins within a similar family of genes exhibiting homology or substantial homology to the suppressing sequence.

### C.4. Other Methods to Inhibit Gene Expression

Yet another means of suppressing gene expression is to insert a polynucleotide into the gene of interest to disrupt transcription or translation of the gene.

Low frequency homologous recombination can be used to target a polynucleotide insert to a gene by flanking the polynucleotide insert with sequences that are substantially similar to the gene to be disrupted. Sequences from the Sequence Listing, fragments thereof, and substantially similar sequence thereto can be used for homologous recombination.

In addition, random insertion of polynucleotides into a host cell genome can also be used to disrupt the gene of interest. Apiroz-Leehan et al., *Trends in Genetics* 13:152 (1997). In this method, screening for clones from a library containing random insertions is preferred to identifying those that have polynucleotides inserted into the gene of interest. Such screening can be performed using probes and/or primers described above based on sequences from the Sequence Listing, fragments thereof, and substantially similar sequence thereto. The screening can also be performed by selecting clones or  $R_1$  plants having a desired phenotype.

### I.D. Methods of Functional Analysis

The constructs described in the methods under I.C. above can be used to determine the function of the polypeptide encoded by the gene that is targeted by the constructs.

Down-regulating the transcription and translation of the targeted gene, the host cell or organisms, such as a plant,

36

may produce phenotypic changes as compared to a wild-type cell or organism. In addition, *in vitro* assays can be used to determine if any biological activity, such as calcium flux, DNA transcription, nucleotide incorporation, etc., are being modulated by the down-regulation of the targeted gene.

Coordinated regulation of sets of genes, e.g., those contributing to a desired polygenic trait, is sometimes necessary to obtain a desired phenotype. SDFs of the invention representing transcription activation and DNA binding domains can be assembled into hybrid transcriptional activators. These hybrid transcriptional activators can be used with their corresponding DNA elements (i.e., those bound by the DNA-binding SDFs) to effect coordinated expression of desired genes (J.J. Schwarz et al., *Mol. Cell. Biol.* 12:266 (1992), A. Martinez et al., *Mol. Gen. Genet.* 261:546 (1999)).

The SDFs of the invention can also be used in the two-hybrid genetic systems to identify networks of protein-protein interactions (L. McAlister-Henn et al., *Methods* 19:330 (1999), J.C. Hu et al., *Methods* 20:80 (2000), M. Golovkin et al., *J. Biol. Chem.* 274:36428 (1999), K. Ichimura et al., *Biochem. Biophys. Res. Comm.* 253:532 (1998)). The SDFs of the invention can also be used in various expression display methods to identify important protein-DNA interactions (e.g. B. Luo et al., *J. Mol. Biol.* 266:479 (1997)).

#### I.E. Promoters

The SDFs of the invention are also useful as structural or regulatory sequences in a construct for modulating the expression of the corresponding gene in a plant or other organism, e.g. a symbiotic bacterium. For example, promoter sequences represented in the Sequence Listing can be useful in directing expression of coding sequences either as constitutive

37

promoters or to direct expression in particular cell types, tissues, or organs or in response to environmental stimuli.

The term "promoter" refers to a region of sequence determinants located upstream or downstream from the start of transcription and which are involved in recognition and binding of RNA polymerase and other proteins to initiate transcription. A "plant promoter" is a promoter capable of initiating transcription in plant cells and can be used to drive expression of a translated portion of an SDF. Such promoters need not be of plant origin. For example, promoters derived from plant viruses, such as the CaMV35S promoter or from *Agrobacterium tumefaciens* such as the T-DNA promoters, can be plant promoters. A typical example of a constitutive promoter of plant origin is the promoter of the cowpea trypsin inhibitor gene. Typical examples of temporal and/or tissue specific promoters of plant origin that can be used with the polynucleotides of the present invention, are: PTA29, a promoter which is capable of driving gene expression specifically in tapetum and only during anther development (Koltonow et al., *Plant Cell* 2:1201 (1990); RCC2 and RCC3, promoters that direct root-specific gene expression in rice (Xu et al., *Plant Mol. Biol.* 27:237 (1995); TobRB27, a root-specific promoter from tobacco (Yamamoto et al., *Plant Cell* 3:371 (1991)).

By "specific promoters" is meant promoters that have a high preference of driving gene expression in the specified tissue and/or at the specified time during the concerned tissue or organ development. By "high preference" is meant at least 3-fold, preferably 5-fold, more preferably at least 10-fold still more preferably at least 20-fold, 50-fold or 100-fold increase in expression in the desired tissue over the expression in any undesired tissue.

A typical example of an inducible promoter, which can be utilized with the polynucleotides of the present invention, is

38

PARSKI, the promoter from the *Arabidopsis* gene encoding a serine-threonine kinase enzyme, and which promoter is induced by dehydration, abscissic acid and sodium chloride (Wang and Goodman, *Plant J.* 8:37 (1995)).

5 With respect to the SDFs of the present invention a promoter is likely to be a relatively small portion of a genomic DNA (gDNA) sequence located in the first 2000 nucleotides upstream from an initial exon identified in a gDNA sequence or initial "ATG" or methionine codon in a corresponding cDNA or mRNA sequence. Such promoters are more likely to be found in the first 1000 nucleotides upstream of an initial ATG or methionine codon of a cDNA sequence corresponding to a gDNA sequence. In particular, the promoter is usually located upstream of the transcription start site.

10 Such a start site is located at the first exon predicted in the OCKHAM-cDNA predictions. In such an instance, the transcription start site is the first nucleotide of the 5' most exon, if the predictions are in the plus (+) strand, or the 3' most if the predictions are in the minus (-) strand. Alternative transcription start sites may be located between the first nucleotide of the 5' most exon (or the 3' most exon in the minus (-) strand) and the initial ATG or methionine codon in the cDNA sequence. The portions of a particular gDNA sequence that function as a promoter in a plant cell will preferably be found to hybridize at medium or high stringency to gDNA sequences presented in the Sequence Listing.

Promoters are generally modular in nature. Short DNA sequences representing binding sites for proteins can be separated from each other by intervening sequences of varying length. For example, within a particular functional module protein binding sites may be constituted by regions of 5 to 60, preferably 10 to 30, more preferably 10 to 20 nucleotides. Within such binding sites, there are typically 2 to 6 nucleotides that specifically contact amino acids of the

39

nucleic acid binding protein. The protein binding sites are usually separated from each other by 10 to several hundred nucleotides, typically by 15 to 150 nucleotides, often by 20 to 50 nucleotides. DNA binding sites in promoter elements often display dyad symmetry in their sequence. Often elements binding several different proteins, and/or a plurality of sites that bind the same protein, will be combined in a region of 100 to 1000 basepairs.

Elements that have transcription regulatory function can be isolated from their corresponding endogenous gene, or the desired sequence can be synthesized, and recombined in constructs to direct expression of a structural gene in a desired tissue-specific, temporal-specific or other desired manner of inducibility or suppression. When hybridizations are performed to identify or isolate elements of a promoter by hybridization to the long sequences presented in the Sequence Listing, conditions should be adjusted to account for the above-described nature of promoters. For example short probes, constituting the element sought, should be used under low temperature and/or high salt conditions. When long probes, which might include several promoter elements are used, low to medium stringency conditions are preferred when hybridizing to promoters across species.

Promoters can consist of a "basal promoter" that functions as a site for assembly of a transcription complex comprising an RNA polymerase, for example RNA polymerase II. A typical transcription complex will include additional factors such as TFIIB, TFIID, and TFIIE. Of these, TFIID appears to be the only one to bind DNA directly. Basal promoters frequently include a "TATA box" element usually located between 20 and 35 nucleotides upstream from the site of initiation of transcription. Basal promoters also sometimes include a "CCAAT box" element (typically a sequence CCAAT) and/or a GGGCG sequence, usually located between 40 and 200 nucleotides,



40

preferably 60 to 120 nucleotides, upstream from the start site of transcription.

The promoter might also contain one or more "enhancers" and/or "suppressors" that function as binding sites for additional transcription factors that have the function of modulating the level of transcription with respect to tissue specificity of transcription, transcriptional responses to particular environmental or nutritional factors, and the like.

If a nucleotide sequence of an SDF, or part of the SDF, functions as a promoter or portion of a promoter, then nucleotide substitutions, insertions or deletions that do not substantially affect the binding of relevant DNA binding proteins would be considered equivalent to the exemplified nucleotide sequence. It is envisioned that there are instances where it is desirable to decrease the binding of relevant DNA binding proteins to "silence" or "down-regulate" a promoter, or conversely to increase the binding of relevant DNA binding proteins to "enhance" or "up-regulate" a promoter. In such instances, polynucleotides representing changes to the nucleotide sequence of the DNA-protein contact region by insertion of additional nucleotides, changes to identity of relevant nucleotides, including use of chemically-modified bases, or deletion of one or more nucleotides are considered encompassed by the present invention.

Promoter function can be assayed by methods known in the art, preferably by measuring activity of a reporter gene operatively linked to the sequence being tested for promoter function. Examples of reporter genes include those encoding luciferase, green fluorescent protein, GUS, neo, cat and bar.

#### I.F. UTRs and Junctions

Polynucleotides comprising untranslated (UTR) sequences and intron/exon junctions are also within the scope of the

41

invention. UTR sequences include introns and 5' or 3' untranslated regions (5' UTRs or 3' UTRs). Portions of the sequences shown in the Sequence Listing can comprise UTRs and introns or intron/exon junctions.

These portions of SDFs, especially UTRs, can have regulatory functions related to, for example, translation rate and mRNA stability. Thus, these portions of SDFs can be isolated for use as elements of gene constructs for expression of polynucleotides encoding desired polypeptides.

Introns of genomic DNA segments might also have regulatory functions. Sometimes promoter elements, especially transcription enhancer or suppressor elements, are found within introns. Also, elements related to stability of heteronuclear RNA and efficiency of transport to the cytoplasm for translation can be found in intron elements. Thus, these segments can also find use as elements of expression vectors intended for use to transform plants.

Just as with promoters, introns and UTR sequences and intron/exon junctions can vary from those shown in the Sequence Listing. Such changes from those sequences preferably will not affect the regulatory activity of the UTRs or intron or intron/exon junction sequences on expression, transcription, or translation. However, in some instances, down-regulation of such activity may be desired to modulate traits or phenotypic or *in vitro* activity.

#### I.G. Coding Sequences

Isolated polynucleotides of the invention can include coding sequences that encode polypeptides comprising an amino acid sequence encoded by a sequence in the Sequence Listing or an amino acid sequence presented in the Sequence Listing.

A nucleotide sequence "encodes" a polypeptide if a cell (or a cell free *in vitro* system) expressing that nucleotide sequence produces a polypeptide having the recited amino acid

42

sequence when the nucleotide sequence is transcribed and the primary transcript is subsequently processed and translated by a host cell (or a cell free *in vitro* system) harboring the nucleic acid. Thus, an isolated nucleic acid that "encodes" a particular amino acid sequence can be a genomic sequence comprising exons and introns or a cDNA sequence that represents the product of splicing thereof. An isolated nucleic acid "encoding an amino acid sequence" also encompasses heteronuclear RNA, which contains sequences that are spliced out during expression, and mRNA, which lacks those sequences.

Coding sequences can be constructed using chemical synthesis techniques or by isolating coding sequences or by modifying such synthesized or isolated coding sequences as described above.

In addition to encoding the polypeptide sequences of the Sequence Listing, which are native to corn or *Arabidopsis*, the isolated polynucleotides can be variant polynucleotides that encode mutants, fragments, and fusions of those native proteins. Such polypeptides are described below in part II.

In variant polynucleotides generally, the number of substitutions, deletions or insertions is preferably less than 20%, more preferably less than 15%; even more preferably less than 10%, 5%, 3% or 1% of the number of nucleotides comprising a particularly exemplified sequence. It is generally expected that non-degenerate nucleotide sequence changes that result in 1 to 10, more preferably 1 to 5 and most preferably 1 to 3 amino acid insertions, deletions or substitutions will not greatly affect the function of an encoded polypeptide. The most preferred embodiments are those wherein 1 to 20, preferably 1 to 10, most preferably 1 to 5 nucleotides are added to, deleted from and/or substituted in the sequences specifically disclosed in the Sequence Listing.

SUBSTITUTE SHEET (RULE 26)

43

Insertions or deletions in polynucleotides intended to be used for encoding a polypeptide should preserve the reading frame. This consideration is not so important in instances when the polynucleotide is intended to be used as a hybridization probe.

## II. Polypeptides

Polypeptides within the scope of the invention include both native proteins as well as mutants, fragments, and fusions thereof. Polypeptides of the invention are those encoded by any of the six reading frames of sequences shown in the Sequence Listing, preferably encoded by the three frames reading in the 5' to 3' direction of the sequences as shown.

Native polypeptides include the proteins encoded by the sequences shown in the Sequence Listing. Such native polypeptides include those encoded by allelic variants.

Variants, including mutants, will exhibit at least 80% sequence identity to those native polypeptides of the Sequence Listing. Sequence identity is used for polypeptides as defined above for polynucleotides. More preferably, the variants will exhibit at least 85% sequence identity; even more preferably, at least 90% sequence identity; more preferably at least 95%, 96%, 97%, 98%, or 99% sequence identity. "Fragments" of polypeptide or "portions" of polypeptides will exhibit similar degrees of identity to the relevant portions of the native polypeptide. Fusions will exhibit similar degrees of identity in that portion of the fusion represented by the variant of the native peptide.

Furthermore, variants will exhibit at least one of the functional properties of the native protein. Such properties include, without limitation, protein interaction, DNA interaction, biological activity, immunological activity, receptor binding, signal transduction, transcription activity,

SUBSTITUTE SHEET (RULE 26)

44

growth factor activity, secondary structure, three-dimensional structure, etc. As to properties related to *in vitro* or *in vivo* activities, the variants preferably exhibit at least 60% of the activity of the native protein; more preferably at least 70%, even more preferably at least 80%, 85%, 90% or 95% of at least one activity of the native protein.

A type of mutant of the native polypeptides comprises amino acid substitutions. "Conservative substitutions" are preferred to maintain the function or activity of the polypeptide. Such substitutions include conservation of charge, polarity, hydrophobicity, size, etc. For example, one or more amino acid residues within the sequence can be substituted with another amino acid of similar polarity that acts as a functional equivalent, for example providing a hydrogen bond in an enzymatic catalysis. Substitutes for an amino acid within an exemplified sequence are preferably made among the members of the class to which the amino acid belongs. For example, the nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan and methionine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine. The positively charged (basic) amino acids include arginine, lysine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

Within the scope of sequence identity described above, a polypeptide of the invention may have additional individual amino acids or amino acid sequences inserted into the polypeptide in the middle thereof and/or at the N-terminal and/or C-terminal ends thereof. Likewise, some of the amino acids or amino acid sequences may be deleted from the polypeptide.

45

#### Antibodies

Isolated polypeptides can be utilized to produce antibodies. Polypeptides of the invention can generally be used, for example, as antigens for raising antibodies by known techniques. The resulting antibodies are useful as reagents for determining the distribution of the antigen protein within the tissues of a plant or within a cell of a plant. The antibodies are also useful for examining the expression level of proteins in various tissues, for example in a wild-type plant or following genetic manipulation of a plant, by methods such as Western blotting.

Antibodies of the present invention, both polyclonal and monoclonal, may be prepared by conventional methods. In general, the polypeptides of the invention are first used to immunize a suitable animal, such as a mouse, rat, rabbit, or goat. Rabbits and goats are preferred for the preparation of polyclonal sera due to the volume of serum obtainable, and the availability of labeled anti-rabbit and anti-goat antibodies as detection reagents. Immunization is generally performed by mixing or emulsifying the protein in saline, preferably in an adjuvant such as Freund's complete adjuvant, and injecting the mixture or emulsion parenterally (generally subcutaneously or intramuscularly). A dose of 50-200 µg/injection is typically sufficient. Immunization is generally boosted 2-6 weeks later with one or more injections of the protein in saline, preferably using Freund's incomplete adjuvant. One may alternatively generate antibodies by *in vitro* immunization using methods known in the art, which for the purposes of this invention is considered equivalent to *in vivo* immunization.

Polyclonal antisera is obtained by bleeding the immunized animal into a glass or plastic container, incubating the blood at 25°C for one hour, followed by incubating the blood at 4°C for 2-18 hours. The serum is recovered by centrifugation

46

(e.g., 1,000xg for 10 minutes). About 20-50 ml per bleed may be obtained from rabbits.

Monoclonal antibodies are prepared using the method of Kohler and Milstein, *Nature* 256: 495 (1975), or modification thereof. Typically, a mouse or rat is immunized as described above. However, rather than bleeding the animal to extract serum, the spleen (and optionally several large lymph nodes) is removed and dissociated into single cells. If desired, the spleen cells can be screened (after removal of nonspecifically adherent cells) by applying a cell suspension to a plate, or well, coated with the protein antigen. B-cells expressing membrane-bound immunoglobulin specific for the antigen bind to the plate, and are not rinsed away with the rest of the suspension. Resulting B-cells, or all dissociated spleen cells, are then induced to fuse with myeloma cells to form hybridomas, and are cultured in a selective medium (e.g., hypoxanthine, aminopterin, thymidine medium, "HAT"). The resulting hybridomas are plated by limiting dilution, and are assayed for the production of antibodies which bind specifically to the immunizing antigen (and which do not bind to unrelated antigens). The selected Mab-secreting hybridomas are then cultured either *in vitro* (e.g., in tissue culture bottles or hollow fiber reactors), or *in vivo* (as ascites in mice).

Other methods for sustaining antibody-producing B-cell clones, such as by EBV transformation, are known.

If desired, the antibodies (whether polyclonal or monoclonal) may be labeled using conventional techniques. Suitable labels include fluorophores, chromophores, radioactive atoms (particularly  $^{32}\text{P}$  and  $^{125}\text{I}$ ), electron-dense reagents, enzymes, and ligands having specific binding partners. Enzymes are typically detected by their activity. For example, horseradish peroxidase is usually detected by its ability to convert 3,3',5,5'-tetramethylbenzidine (TNB) to a blue pigment,

47

quantifiable with a spectrophotometer. "Specific binding partner" refers to a protein capable of binding a ligand molecule with high specificity, as for example in the case of an antigen and a monoclonal antibody specific therefor. Other specific binding partners include biotin and avidin or streptavidin, IgG and protein A, and the numerous receptor-ligand couples known in the art. It should be understood that the above description is not meant to categorize the various labels into distinct modes. For example,  $^{125}\text{I}$  may serve as a radioactive label or as an electron-dense reagent. HRP may serve as an enzyme or as an antigen for a Mab. Further one may combine various labels for desired effect. For example, Mabs and avidin also require labels in the practice of this invention: thus, one might label a Mab with biotin, and detect its presence with avidin labeled with  $^{125}\text{I}$ , or with an anti-biotin Mab labeled with HRP. Other permutations and possibilities will be readily apparent to those of ordinary skill in the art, and are considered as equivalents within the scope of the instant invention.

#### 20 In Vitro Applications of Polypeptides

Some polypeptides of the invention will have enzymatic activities that are useful *in vitro*. For example, the soybean trypsin inhibitor (Kunitz) family is one of the numerous families of proteinase inhibitors. It comprises plant proteins which have inhibitory activity against serine proteinases from the trypsin and subtilisin families, thiol proteinases and aspartic proteinases. Thus, these peptides find *in vitro* use in protein purification protocols and perhaps in therapeutic settings requiring topical application of protease inhibitors.

Delta-aminolevulinic acid dehydratase (EC 4.2.1.24) (ALAD) catalyzes the second step in the biosynthesis of heme, the condensation of two molecules of 5-aminolevulinate to

48

form porphobilinogen. Thus, ALAD proteins can be used as catalysts in synthesis of heme derivatives. Enzymes of biosynthetic pathways generally can be used as catalysts for in vitro synthesis of the compounds representing products of the pathway.

5

Polypeptides encoded by SDFs of the invention can be engineered to provide purification reagents to identify and purify additional polypeptides that bind to them. This allows one to identify proteins that function as multimers or elucidate signal transduction or metabolic pathways. In the case of DNA binding proteins, the polypeptide can be used in a similar manner to identify the DNA determinants of specific binding (S. Pierrou et al., *Anal. Biochem.* 229:99 (1995), S. Chusacutanachai et al., *J. Biol. Chem.* 274:23591 (1999), Q. Lin et al., *J. Biol. Chem.* 272:27274 (1997)).

15

#### II.A. MUTANTS, FRAGMENTS, AND FUSIONS

Generally, mutants, fragments, or fusions of the polypeptides encoded by the maximum length sequence (MLS) can exhibit at least one of the activities of the identified domains and/or related polypeptides described in Sections (C) and (D) of TABLE 1 corresponding to the MLS of interest.

20

##### II.A.(1) Mutants

A type of mutant of the native polypeptides comprises amino acid substitutions. "Conservative substitutions", described above (see II.), are preferred to maintain the function or activity of the polypeptide.

25

Such substitutions include conservation of charge, polarity, hydrophobicity, size, etc. For example, one or more amino acid residues within the sequence can be substituted with another amino acid of similar polarity that acts as a functional equivalent, for example providing a hydrogen bond in an enzymatic catalysis. Substitutes for an

30

49

amino acid within an exemplified sequence are preferably made among the members of the class to which the amino acid belongs. For example, the nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan and methionine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine. The positively charged (basic) amino acids include arginine, lysine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

5

10

Within the scope of sequence identity described above, a polypeptide of the invention may have additional individual amino acids or amino acid sequences inserted into the polypeptide in the middle thereof and/or at the N-terminal and/or C-terminal ends thereof. Likewise, some of the amino acids or amino acid sequences may be deleted from the polypeptide. Amino acid substitutions may also be made in the sequences; conservative substitutions being preferred.

15

One preferred class of mutants are those that comprise (1) the domain of a MLS encoded polypeptide and/or (2) residues conserved between the MLS encoded polypeptide and related polypeptides of the MLS. For this class of mutants, the MLS encoded polypeptide sequence is changed by insertion, deletion, or substitution at positions flanking the domain and/or conserved residues.

25

Another class of mutants includes those that comprise a MLS encoded polypeptide sequence that is changed in the domain or conserved residues by a conservative substitution.

30

Yet another class of mutants includes those that lack one of the in vitro activities, or structural features of the MLS encoded polypeptides. One example is dominant negative mutants. Such a mutant may comprise an MLS encoded

50

polypeptide sequence with non-conservative changes in a particular domain or group of conserved residues.

#### II.A.(2) FRAGMENTS

Fragments of particular interest are those that comprise a domain identified for a polypeptide encoded by an MLS of the instant invention and mutants thereof. Also, fragments that comprise at least one region of residues conserved between an MLS encoded polypeptide and its related polypeptides are of great interest. Fragments are sometimes useful as dominant negative mutations.

#### II.A.(3) FUSIONS

Of interest are chimeras comprising (1) a fragment of the MLS encoded polypeptide or mutants thereof of interest and (2) a fragment of a polypeptide comprising the same domain. For example, an AP2 helix encoded by a MLS of the invention fused to second AP2 helix from ANT protein, which comprises two AP2 helices. The present invention also encompasses fusions of MLS encoded polypeptides, mutants, or fragments thereof fused with related proteins or fragments thereof.

#### DEFINITION OF DOMAINS

The polypeptides of the invention may possess identifying domains as shown in TABLE 1. Domains are fingerprints or signatures that can be used to characterize protein families and/or motifs. Such fingerprints or signatures can comprise conserved (1) primary sequence, (2) secondary structure, and/or (3) three-dimensional conformation. Generally, each domain has been associated with either a family of proteins or a motif. Typically, these families and/or motifs have been correlated with specific *in-vitro* and/or *in-vivo* activities. A domain can be any length, including the entirety of the sequence of a

51

Detailed descriptions of the domains, associated families and motifs, and correlated activities of the polypeptides of the instant invention are described below. Usually, the polypeptides with designated domain(s) can exhibit at least one activity that is exhibited by any polypeptide that comprises the same domain(s).

Specific domains within the MLS encoded polypeptides are indicated by the reference TABLE 1. In addition, the domains within the MLS encoded polypeptide can be defined by the region that exhibits at least 70% sequence identity with the consensus sequences listed in the detailed description below of each of the domains.

The majority of the protein domain descriptions given below are obtained from Prosite, (<http://www.expasy.ch/prosite/>), and Pfam, (<http://pfam.wustl.edu/browse.shtml>).

52

# 1. (AAA) AAA-protein family signature

A large family of ATPases has been described [1 to 5] whose key feature is that they share a conserved region of about 220 amino acids that contains an ATP-binding site. This family

5 is now called AAA, for 'A'TPases 'A'ssociated with diverse cellular 'A'ctivities. The proteins that belong to this family either contain one or two AAA domains. Proteins containing two AAA domains:

10 - Mammalian and drosophila NSF (N-ethylmaleimide-sensitive fusion protein) and the fungal homolog, SEC18. These proteins are involved in intracellular transport between the endoplasmic reticulum and Golgi, as well as between different Golgi cisternae.

- Mammalian transitional endoplasmic reticulum ATPase

15 (previously known as p97 or VCP) which is involved in the transfer of membranes from the endoplasmic reticulum to the golgi apparatus. This protein forms a ring-shaped

homooligomer composed of six subunits. The yeast homolog is CDC48 and it may play a role in spindle pole proliferation.

20 - Yeast protein PAS1, essential for peroxisome assembly and the related protein PAS1 from *Pichia pastoris*.

- Yeast protein AFG2.

- *Sulfolobus acidocaldarius* protein SAV and *Halobacterium salinarum* cdch which may be part of a transduction pathway connecting light to cell division.

25 Proteins containing a single AAA domain:

- *Escherichia coli* and other bacteria ftsh (or hflB) protein. Ftsh is an ATP-dependent zinc metallopeptidase that seems to degrade the heat-shock sigma-32 factor.

30 It is an integral membrane protein with a large cytoplasmic C-terminal domain that contain both the AAA and the protease domains.

53

- Yeast protein YME1, a protein important for maintaining the integrity of the mitochondrial compartment. YME1 is also a zinc-dependent protease.

5 - Yeast protein AFG3 (or YTA10). This protein also seems to contain a AAA domain followed by a zinc-dependent protease domain.

Subunits from the regulatory complex of the 26S

proteasome [6] which is involved in the ATP-dependent degradation of ubiquitinated proteins:

10 a) Mammalian subunit 4 and homologs in other higher eukaryotes, in yeast (gene YTA5) and fission yeast (gene mts2).

b) Mammalian subunit 6 (TBP7) and homologs in other higher eukaryotes and in yeast (gene YTA2).

15 c) Mammalian subunit 7 (MSS1) and homologs in other higher eukaryotes and in yeast (gene CIM5 or YTA3).

d) Mammalian subunit 8 (P45) and homologs in other higher eukaryotes and in yeast (SUG1 or CIM3 or TBY1) and fission yeast (gene let1).

20 Other probable subunits such as human TBP1 which seems to influences HIV gene expression by interacting with the virus tat transactivator protein and yeast YTA1 and YTA6.

- Yeast protein BCS1, a mitochondrial protein essential for the expression of the Rieske iron-sulfur protein.

25 - Yeast protein MSP1, a protein involved in intramitochondrial sorting of proteins.

- Yeast protein PAS8, and the corresponding proteins PAS5 from *Pichia pastoris* and PAY4 from *Yarrowia lipolytica*.

30 - Mouse protein SKD1 and its fission yeast homolog (SPAC2G11.06).

- *Caenorhabditis elegans* meiotic spindle formation protein mei-1.

- Yeast protein SAPI.

54

- Yeast protein YTA7.
  - Mycobacterium leprae hypothetical protein A2126A.
- It is proposed that, in general, the AAA domains in these proteins act as ATP- dependent protein clamps [5]. In addition to the ATP-binding 'A' and 'B' motifs, which are located in the N-terminal half of this domain, there is a highly conserved region located in the central part of the domain which was used to develop a signature pattern.

10 Consensus pattern: [LIVMT]-x-[LIVMT]-[LIVMF]-x-[GATMC]-[ST]-[NS]-x(4)-[LIVM]-D-x-A-[LIFA]-x-R

[1] Froehlich K.-U., Fries H.W., Ruediger M., Erdmann R., Botstein D., Mecke D. J. Cell Biol. 114:443-453(1991).

15 [2] Erdmann R., Wiebel F.F., Flessau A., Rytko J., Beyer A., Froehlich K.-U., Kunau W.-H. Cell 64:499-510(1991).

[3] Peters J.-M., Walsh M.J., Franke W.W. EMBO J. 9:1757-1767(1990).

20 [4] Kunau W.-H., Beyer A., Goette K., Marzloch M., Saidowsky J., Skaletz-Rorowski A., Wiebel F.F. Biochimie 75:209-224(1993).

[5] Confalonieri F., Duguët M. BioEssays 17:639-650(1995).

6] Hilt W., Wolf D.H. Trends Biochem. Sci. 21:96-102(1996).

## 25 2. Aminotransferases class-IV signature

Aminotransferases share certain mechanistic features with other pyridoxal-phosphate dependent enzymes, such as the covalent binding of the pyridoxal-phosphate group to a lysine residue. On the basis of sequence similarity, these various enzymes can be grouped [1,2] into subfamilies. One of these, called class-IV, currently consists of the following enzymes:

- Branched-chain amino-acid aminotransferase (EC 2.6.1.42) (transaminase B), a bacterial (gene *ilvE*) and eukaryotic enzyme which catalyzes the reversible

55

transfer of an amino group from 4-methyl-2-oxopentanoate to glutamate, to form leucine and 2-oxoglutarate.

- D-alanine aminotransferase (EC 2.6.1.21). A bacterial enzyme which catalyzes the transfer of the amino group from D-alanine (and other D-amino acids) to 2-oxoglutarate, to form pyruvate and D-aspartate.
- 4-amino-4-deoxychorismate (ADC) lyase (gene *pabC*). A bacterial enzyme that converts ADC into 4-aminobenzoate (PABA) and pyruvate.

10 The above enzymes are proteins of about 270 to 415 amino-acid residues that share a few regions of sequence similarity. Surprisingly, the best-conserved region does not include the lysine residue to which the pyridoxal-phosphate group is known to be attached, in *ilvE*. The region that has been selected as a signature pattern is located some 40 residues at the C-terminus side of the PLP-lysine

Consensus pattern: E-x-[STAGCI]-x(2)-N-[LIVMFAC]-[FY]-x(6,12)-[LIVMF]-x-T-x(6,8)-[LIVM]-x-[GS]-[LIVM]-x-[KR]-

[1] Green J.M., Merkel W.K., Nichols B.P. J. Bacteriol. 174:5317-5323(1992).

[2] Bairoch A. Unpublished observations (1992).

## 25 3. Bacterial mutT domain signature

The bacterial mutT protein is involved in the GO system [1] responsible for removing an oxidatively damaged form of guanine (8-hydroxyguanine or 7,8-dihydro-8-oxoguanine) from DNA and the nucleotide pool. 8-oxo-dGTP is inserted opposite to dA and dC residues of template DNA with almost equal efficiency thus leading to A.T to G.C transversions. MutT specifically degrades 8-oxo-dGTP to the monophosphate with the concomitant release of pyrophosphate. MutT is a small



56

protein of about 12 to 15 Kd. It has been shown [2,3] that a region of about 40 amino acid residues, which is found in the N-terminal part of mutT, can also be found in a variety of other prokaryotic, viral, and eukaryotic proteins. These proteins are:

- Streptomyces pneumoniae mutX.
- A mutT homolog from plasmid pSAM2 of Streptomyces ambofaciens.
- Bartonella bacilliformis invasion protein A (gene invA).
- Escherichia coli dATP pyrophosphohydrolase.
- Protein D250 from African swine fever viruses.
- Proteins D9 and D10 from a variety of poxviruses.
- Mammalian 7,8-dihydro-8-oxoguanine triphosphatase (EC 3.1.6.-) [4].
- Mammalian diadenosine 5',5'''-P<sub>1</sub>P<sub>4</sub>-tetrakisphosphate asymmetrical hydrolase (Ap4Aase) (EC 3.6.1.17) [5], which cleaves A-5'-PPP-5'A to yield AMP and ATP.
- A protein encoded on the antisense RNA of the basic fibroblast growth factor gene in higher vertebrates.
- Yeast protein YSA1.
- Escherichia coli hypothetical protein yfaO.
- Escherichia coli hypothetical protein ygdU and HI0901, the corresponding Haemophilus influenzae protein.
- Escherichia coli hypothetical protein yJad and HI0432, the corresponding Haemophilus influenzae protein.
- Escherichia coli hypothetical protein yrfE.
- Bacillus subtilis hypothetical protein yqKG.
- Bacillus subtilis hypothetical protein yzgd.
- Yeast hypothetical protein YGL067w.

It is proposed [2] that the conserved domain could be involved in the active center of a family of pyrophosphate-releasing NTPases. As a signature pattern the core region of

57

the domain was selected; it contains four conserved glutamate residues.

Consensus pattern: G-x(5)-E-x(4)-[STAGC]-[LIVMAC]-x-R-E-

- 5 [LIVMFT]-x-E-E-
- [1] Michaels M.L., Miller J.H. J. Bacteriol. 174:6321-6325(1992).
- [2] Koonin E.V. Nucleic Acids Res. 21:4847-4847(1993).
- 10 [3] Mejean V., Salles C., Bullions M.J., Bessman M.J., Claverys J.-P. Mol. Microbiol. 11:323-330(1994).
- [4] Sakumi K., Furuichi M., Tsuzuki T., Kakuma T., Kawabata S., Maki H., Sekiguchi M. J. Biol. Chem. 268:23524-23530(1993).
- 15 [5] Thorne N.M.H., Hankin S., Wilkinson M.C., Nunez C., Barracough R., McLennan A.G. Biochem. J. 311:717-721(1995).

#### 4. Cystatin domain

This is a very diverse family. Attempts to define separate subfamilies have failed. Typically, either the N-terminal or C-terminal end is very divergent. But splitting into two domains would make very short families. Cathelicidins are related to this family but have not been included. Number of members: 147

- 25 Inhibitors of cysteine proteases [1,2,3], which are found in the tissues and body fluids of animals, in the larva of the worm Onchocerca volvulus [4], as well as in plants, can be grouped into three distinct but related families:
  - Type 1 cystatins (or stefins), molecules of about 100 amino acid residues with neither disulfide bonds nor carbohydrate groups.
  - Type 2 cystatins, molecules of about 115 amino acid residues which contain one or two disulfide loops near their C-terminus.
- 30

58

- Kininogens, which are multifunctional plasma glycoproteins.

They are the precursor of the active peptide bradykinin and play a role in blood coagulation by helping to position optimally prekallikrein and factor XI next to factor XII.

They are also inhibitors of cysteine proteases. Structurally, kininogens are made of three contiguous type-2 cystatin domains, followed by an additional domain (of variable length) which contains the sequence of bradykinin. The first of the three cystatin domains seems to have lost its inhibitory activity.

In all these inhibitors, there is a conserved region of five residues which has been proposed to be important for the binding to the cysteine proteases. The consensus pattern starts one residue before this conserved region.

-Consensus pattern: [GSTEQKRV]-Q-[LIVT]-[VAF]-[SAGQ]-G-x-[LIVMNK]-x(2)-[LIVMFY]-x-[LIVMFYA]-[DENQKRHSIV]

- [1] Barrett A.J. Trends Biochem. Sci. 12:193-196(1987).
- [2] Rawlings N.D., Barrett A.J. J. Mol. Evol. 30:60-71(1990).
- [3] Turk V., Bode W. FEBS Lett. 285:213-219(1991).
- [4] Lustigman S., Brotman B., Huima T., Prince A.M. Mol. Biochem. Parasitol. 45:65-76(1991).

#### 5. Dehydrins signatures

A number of proteins are produced by plants that experience water-stress. Water-stress takes place when the water available to a plant falls below a critical level. The plant hormone abscisic acid (ABA) appears to modulate the response of plant to water-stress. Proteins that are expressed during water-stress are called dehydrins [1,2] or LEA group 2 proteins [3]. The proteins that belong to this family are listed below.

59

- Arabidopsis thaliana XERO 1, XERO 2 (LTI30), RAB18, ERD10 (LTI45) ERD14 and COR47.

- Barley dehydrins B8, B9, B17, and B18.

- Cotton LEA protein D-11.

- Craterostigma plantagineum desiccation-related proteins A and B.

- Maize dehydrin M3 (RAB-17).

- Pea dehydrins DHN1, DHN2, and DHN3.

- Radish LEA protein.

- Rice proteins RAB 16B, 16C, 16D, RAB21, and RAB25.

- Tomato TASI4.

- Wheat dehydrin RAB 15 and cold-shock protein cor410, cs66 and cs120.

Dehydrins share a number of structural features.

One of the most notable features is the presence, in their central region, of a continuous run of five to nine serines followed by a cluster of charged residues.

Such a region has been found in all known dehydrins so far with the exception of pea dehydrins. A second

conserved feature is the presence of two copies of alysine-rich octapeptide; the first copy is located just after the cluster of charged residues that follows the poly-serine region and the second copy is found at the C-terminal extremity. Signature patterns for both regions were derived.

Consensus pattern: S(5)-[DE]-x-[DE]-G-x(1,2)-G-x(0,1)-[KR] (4)  
Consensus pattern: [KR]-[LIM]-K-K-[DE]-K-[LIM]-P-G-

- [1] Close T.J., Kortt A.A., Chandler P.M. Plant Mol. Biol. 13:95-108(1989).
- [2] Robertson M., Chandler P.M. Plant Mol. Biol. 19:1031-1044(1992).

[3] Dure L. III, Crouch M., Harada J., Ho T.-H. D., Mundy J., Quatrano R., Thomas T., Sung Z.R. Plant Mol. Biol. 12:475-486(1989).

5 6. D-isomer specific 2-hydroxyacid dehydrogenases (2 Hacid DH)  
This Pfam covers the Formate dehydrogenase, D-glycerate dehydrogenase and D-lactate dehydrogenase families in SCOP. A number of NAD-dependent 2-hydroxyacid dehydrogenases which seem to be specific for the D-isomer of their substrate have been shown [1,2,3,4] to be functionally and structurally related. These enzymes are listed below.

- D-lactate dehydrogenase (EC 1.1.1.28), a bacterial enzyme which catalyzes the reduction of D-lactate to pyruvate.

15 - D-glycerate dehydrogenase (EC 1.1.1.29) (NADH-dependent hydroxypyruvate reductase), a plant leaf peroxisomal enzyme that catalyzes the reduction of hydroxypyruvate to glycerate. This reaction is part of the glycolate pathway of photorespiration.

- D-glycerate dehydrogenase from the bacteria *Hyphomicrobium methylovorum* and *Methylobacterium extorquens*.

25 - 3-phosphoglycerate dehydrogenase (EC 1.1.1.95), a bacterial enzyme that catalyzes the oxidation of D-3-phosphoglycerate to 3-phosphohydroxypyruvate. This reaction is the first committed step in the 'phosphorylated' pathway of serine biosynthesis.

- Erythronate-4-phosphate dehydrogenase (EC 1.1.1.-) (gene *pdxB*), a bacterial enzyme involved in the biosynthesis of pyridoxine (vitamin B6).

30 - D-2-hydroxyisocaproate dehydrogenase (EC 1.1.1.-) (*hicDH*), a bacterial enzyme that catalyzes the reversible and stereospecific interconversion between

2-ketocarboxylic acids and D-2-hydroxy-carboxylic acids.

- Formate dehydrogenase (EC 1.2.1.2) (FDH) from the bacteria *Pseudomonas* sp. 101 and various fungi [5].
- 5 - Vancomycin resistance protein *vanH* from *Enterococcus faecium*; this protein is a D-specific alpha-keto acid dehydrogenase involved in the formation of a peptidoglycan which does not terminate by D-alanine thus preventing vancomycin binding.
- 10 - *Escherichia coli* hypothetical protein *ycdW*.
- *Escherichia coli* hypothetical protein *viaE*.
- *Haemophilus influenzae* hypothetical protein *HI1556*.
- Yeast hypothetical protein *YER081w*.
- Yeast hypothetical protein *YIL074w*.

15 All these enzymes have similar enzymatic activities and are structurally related. Three of the most conserved regions of these proteins have been selected to develop patterns. The first pattern is based on a glycine-rich region located in the central section of these enzymes; this region probably corresponds to the NAD-binding domain. The two other patterns contain a number of conserved charged residues, some of which may play a role in the catalytic mechanism.

25 -Consensus pattern: [LIVMA]-[AG]-[IVT]-[LIVMFY]-[AG]-x-G-[NHRQGSAC]-[LIV]-G-x(13,14)-[LIVEMT]-x(2)-[FYWCTH]-[DNSTK]-  
-Consensus pattern: [LIVMEFWA]-[LIVFYWC]-x(2)-[SAC]-[DNQHR]-[IVFA]-[LIVF]-x-[LIVF]-[HNI]-x-P-x(4)-[STN]-x(2)-[LIVMF]-x-[GSDN]

30 -Consensus pattern: [LMFATC]-[KPQ]-x-[GSTDN]-x-[LIVMEFYWR]-[LIVMEFYW](2)-N-x-[STAGC]-R-[GP]-x-[LIVH]-[LIVMC]-[DNV]

[1] Grant G.A. Biochem. Biophys. Res. Commun. 165:1371-1374 (1989).

- [2] Kochhar S., Hunziker P., Leong-Morgenthaler P.M., Hottinger H. Biochem. Biophys. Res. Commun. 184:60-66(1992).  
 [3] Ohta T., Taguchi H. J. Biol. Chem. 266:12588-12594(1991).  
 [4] Goldberg J.D., Yoshida T., Brick P. J. Mol. Biol. 236:1123-1140(1994).  
 [5] Popov V.O., Lamzin V.S. Biochem. J. 301:625-643(1994).

#### 7. dnaJ domains signatures and profile

The prokaryotic heat shock protein dnaJ interacts with the chaperone hsp70-like dnaK protein [1]. Structurally, the dnaJ protein consists of an N-terminal conserved domain (called 'J' domain) of about 70 amino acids, a glycine-rich region ('G' domain') of about 30 residues, a central domain containing four repeats of a CXXCXGXG motif ('CRR' domain) and a C-terminal region of 120 to 170 residues. Such a structure is shown in the following schematic representation:

```

+-----+-----+-----+-----+-----+-----+-----+
-----+ | N-terminal | | Gly-R | | CXXCXGXG | C-
terminal | +-----+-----+-----+-----+-----+
-----+

```

It has been shown [2] that the 'J' domain as well as the 'CRR' domain are also found in other prokaryotic and eukaryotic proteins which are listed below.

#### a) Proteins containing both a 'J' and a 'CRR' domain:

- Yeast protein MAS5/YDJ1 which seems to be involved in mitochondrial protein import.
- Yeast protein MDJ1, involved in mitochondrial biogenesis and protein folding.
- Yeast protein SCJ1, involved in protein sorting.
- Yeast protein XDJ1.
- Plants dnaJ homologs (from leek and cucumber).
- Human HDJ2, a dnaJ homolog of unknown function.
- Yeast hypothetical protein YNL077w.

#### b) Proteins containing a 'J' domain without a 'CRR' domain:

- Rhizobium fredii nolC, a protein involved in cultivar-specific nodulation of soybean.
- Escherichia coli cbpA [3], a protein that binds curved DNA.
- Yeast protein SEC63/NPL1, important for protein assembly into the endoplasmic reticulum and the nucleus.

- Yeast protein SISI, required for nuclear migration during mitosis.

- Yeast protein CAJ1.

- Yeast hypothetical protein YFR041c.

- Yeast hypothetical protein YIR004w.

- Yeast hypothetical protein YJL162c.

- Plasmodium falciparum ring-infected erythrocyte

surface antigen (RESA). RESA, whose function is not known, is associated with the membrane skeleton of newly invaded erythrocytes.

- Human HDJ1.

- Human HSJ1, a neuronal protein.

- Drosophila cysteine-string protein (csp).

A signature pattern for the 'J' domain was developed, based on conserved positions in the C-terminal half of this domain. A pattern for the 'CRR' domain, based on the first two copies of that motif was also developed. A profile for the 'J' domain was also developed.

Consensus pattern: [FY]-x(2)-[LIVMA]-x(3)-[FYWHNT]-[DENQSA]-x-L-x-[DN]-x(3)-[KR]-x(2)-[FYI]-  
 Consensus pattern: C-[DEGSTHKR]-x-C-x-G-x-[CK]-[AGSDM]-x(2)-[GSNKR]-x(4,6)-C-x(2,3)-C-x-G-x-G-

[1] Cyr D.M., Langer T., Douglas M.G. Trends Biochem. Sci. 19:176-181(1994).

64

[2] Bork P., Sander C., Valencia A., Bukau B. Trends Biochem. Sci. 17:129-129(1992).

[3] Ueguchi C., Kaneda M., Yamada H., Mizuno T. Proc. Natl. Acad. Sci. U.S.A. 91:1054-1058(1994).

5

## 8. Domain of unknown function

9. Gamma-thionins family signature

The following small plant proteins are evolutionary

10 related:

- Gamma-thionins from wheat endosperm (gamma-purothionins) and barley (gamma-hordothionins) which are toxic to animal cells and inhibit protein synthesis in cell free systems [1].

15 - A flower-specific thionin (EST) from tobacco [2].

- Antifungal proteins (AFP) from the seeds of Brassicaceae species such as radish, mustard, turnip and *Arabidopsis thaliana* [3].

- Inhibitors of insect alpha-amylases from sorghum [4].

20 - Probable protease inhibitor P322 from potato.

- A germination-related protein from cowpea [5].

- Anther-specific protein SF18 from sunflower [6]. SF18 is a protein that contains a gamma-thionin domain at its N-terminus and a proline-rich C-terminal domain.

25 - Soybean sulfur-rich protein SE60 [7].

- *Vicia faba* antibacterial peptides fabatin-1 and -2.

In their mature form, these proteins generally consist of about 45 to 50 amino-acid residues. As shown in the following schematic representation, these peptides contain eight conserved cysteines involved in disulfide bonds.

[illegible]

— + — — —

xxxxxxxxxxxxxxCxxxxCxxxxx

[illegible]

35 +

59

'C': conserved cysteine involved in a disulfide bond.  
'\*': position of the pattern.

Consensus pattern: [KRG]-x-C-x(3)-[SV]-x(2)-[FYWH]-x-[GF]-x-  
5 C-x(5)-C-x(3)-C [The four C's are involved in disulfide  
bonds]-

[1] Bruix M., Jimenez M.A., Santoro J., Gonzalez C., Collila F.J., Mendez E., Rico M. *Biochemistry* 32:715-724 (1993).

10 [2] Gu Q., Kawata E.E., Morse M.-J., Wu H.-M., Cheung A.Y.  
Mol. Gen. Genet. 234:89-96(1992).

[3] Terras F.R.G., Torrekens S., van Leuven F., Osborn R.W., Vanderleyden J., Cammue B.P.A., Broekaert W.F. *FEBS Lett.* 316:233-240(1993).

15 [4] Bloch C. Jr., Richardson M. FEBS Lett. 279:101-104(1991).

[5] Ishibashi N., Yamauchi D., Miniamikawa T. Plant Mol. Biol. 15:59-64(1990).

[7] Choi Y., Choi Y.D., Lee J.S. Plant Physiol. 101:699-700 (1993).

20

10. haloacid dehalogenase-like hydrolase

This family is structurally different from the alpha/beta hydrolase family (abhydrolase). This family includes L-2-haloacid dehalogenase, epoxide hydrolases and phosphatases.

25 The structure of the family consists of two domains. One is an inserted four helix bundle, which is the least well conserved region of the alignment, between residues 16 and 96 of Swiss:P24069. The rest of the fold is composed of the core alpha/beta domain.

30 [1] Hisano T, Hata Y, Fujii T, Liu JQ, Kurihara T, Esaki N, Soda K. *J Biol Chem* 1996; 271:20322-20330.

## 11. Helix-turn-helix (HTH)

66

This large family of DNA binding helix-turn helix proteins includes Cro Swiss:P03036 and CI Swiss:P03034.

12. Heme-binding domain in cytochrome b5 and oxidoreductases

5 (heme\_1)

Cytochrome b5 is a membrane-bound hemo protein which acts as an electron carrier for several membrane-bound oxygenases [1]. There are two homologous forms of b5, one found in microsomes and one found in the outer membrane of mitochondria. Two conserved histidine residues serve as axial ligands for the heme group. The structure of a number of oxidoreductases consists of the juxtaposition of a heme-binding domain homologous to that of b5 and either a flavodehydrogenase or a molybdopterin domain. These enzymes are:

- Lactate dehydrogenase (EC 1.1.2.3) [2], an enzyme that consists of a flavodehydrogenase domain and a heme-binding domain called cytochrome b2.
  - Nitrate reductase (EC 1.6.6.1), a key enzyme involved in the first step of nitrate assimilation in plants, fungi and bacteria [3,4]. Consists of a molybdopterin domain (see <PDOC00484>), a heme-binding domain called cytochrome b557, as well as a cytochrome reductase domain.
  - Sulfite oxidase (EC 1.8.3.1) [5], which catalyzes the terminal reaction in the oxidative degradation of sulfur-containing amino acids. Also consists of a molybdopterin domain and a heme-binding domain.
- This family of proteins also includes:
- TU-36B, a *Drosophila* muscle protein of unknown function [6].
  - Fission yeast hypothetical protein SpAC1F12.10c.
  - Yeast hypothetical protein YMR073c.
  - Yeast hypothetical protein YMR272c.

SUBSTITUTE SHEET (RULE 26)

67

A segment was used which includes the first of the two histidine heme ligands, as a signature pattern for the heme-binding domain of cytochrome b5 family.

5 Consensus pattern: [FY]-[LIVMK]-x(2)-H-P-[GA]-G [H is a heme axial ligand]-

[1] Ozols J. Biochim. Biophys. Acta 997:121-130(1989).

[2] Guillard B. EMBO J. 4:3265-3272(1985).

10 [3] Calza R., Huttner E., Vincentz M., Rouze P., Galangau F., Vaucheret H., Cherel I., Meyer C., Kronenberger J., Caboche M. Mol. Gen. Genet. 209:552-562(1987).

[4] Crawford N.M., Smith M., Bellissimo D., Davis R.W. Proc. Natl. Acad. Sci. U.S.A. 85:5006-5010(1988).

15 [5] Guillard B., Lederer F. Eur. J. Biochem. 100:441-453(1979).

[6] Levin R.J., Boychuk P.L., Croniger C.M., Kazzaz J.A., Rozek C.E. Nucleic Acids Res. 17:6349-6367(1989).

13. KH domain

20 KH motifs probably bind RNA directly. Auto antibodies to Nova, a KH domain protein, cause paraneoplastic opsoclonus ataxia.

[1] Burd CG, Dreyfuss G, Science 1994;265:615-621.

[2] Musco G, Stier G, Joseph C, Castiglione Morelli MA,

25 Nilges M, Gibson TJ, Pastore A, Cell 1996;85:237-245.

14. MAPEG family (aka: FLAP/GST2/LTC4S family signature)

The following mammalian proteins are evolutionary related [1]:

- 30 - Leukotriene C4 synthase (EC 2.5.1.37) (gene LTC4S), an enzyme that catalyzes the production of LTC4 from LTA4.

SUBSTITUTE SHEET (RULE 26)

68

- Mitochondrial glutathione S-transferase II (EC 2.5.1.18) (GST-II) (gene GST2), an enzyme that can also produce LTC4 from LTA4.
- 5-lipoxygenase activating protein (gene FLAP), a protein that seems to be required for the activation of 5-lipoxygenase.

These are proteins of 150 to 160 residues that contain three transmembrane segments. As a signature pattern, a conserved region between the first and second transmembrane domains was selected.

Consensus pattern: G-x(3)-F-E-R-V-[FY]-x-A-[NQ]-x-N-C

- [1] Jakobsson P.-J., Mancini J.A., Ford-Hutchinson A.W. J. Biol. Chem. 271:22203-22210 (1996).

#### 15. Pathogenesis-related protein Bet v I family signature

A number of plant proteins, which all seem to be involved in pathogen defense response, are structurally related [1,2,3]. These proteins are:

- Bet v I, the major pollen allergen from white birch. Bet v I is the main cause of type I allergic reactions in Europe, North America and USSR.
- Aln g I, the major pollen allergen from alder.
- Api G I, the major allergen from celery.
- Car b I, the major pollen allergen from hornbeam.
- Cor a I, the major pollen allergen from hazel.
- Mal d I, the major pollen allergen from apple.
- Asparagus wound-induced protein AopRI.
- Kidney bean pathogenesis-related proteins 1 and 2.
- Parsley pathogenesis-related proteins PR1-1 and PR1-3.
- Pea disease resistance response proteins pI49, pI176 and DRG49-C.
- Pea abscisic acid-responsive proteins ABR17 and ABR18.

SUBSTITUTE SHEET (RULE 26)

69

- Potato pathogenesis-related proteins STH-2 and STH-21.
- Soybean stress-induced protein SAM22.

These proteins are thought to be intracellularly

located. They contain from 155 to 160 amino acid residues. As a signature pattern, a conserved region located in the third quarter of these proteins has been selected

Consensus pattern: G-x(2)-[LIVMF]-x(4)-E-x(2)-[CSTAEN]-x(8,9)-[GND]-G-[GS]-[CS]-x(2)-K-x(4)-[FY]-

- [1] Breiteneder H., Pettenburger K., Bito A., Valentia R., Kraft D., Rumpold H., Scheiner O., Breitenbach M. EMBO J. 8:1935-1938 (1989).

[2] Crowell D., John M.E., Russell D., Amasino R.M. Plant Mol. Biol. 18:459-466 (1992).

- [3] Warner S.A.J., Scott R., Draper J. Plant Mol. Biol. 19:555-561 (1992).

#### 16. Photosystem I psaG / psak (PSI PSAK) proteins signature

Photosystem I (PSI) [1] is an integral membrane protein

complex that uses light energy to mediate electron transfer from plastocyanin to ferredoxin. It is found in the chloroplasts of plants and cyanobacteria. PSI is composed of at least 14 different subunits, two of which PSI-G (gene psaG) and PSI-K (gene psak) are small hydrophobic proteins of about 7 to 9 Kd and evolutionary related [2]. Both seem to contain two transmembrane regions. Cyanobacteria seem to encode only for PSI-K.

As a signature pattern, the best-conserved region was selected which seems to correspond to the second transmembrane region.

-Consensus pattern: [GT]-F-x-[LIVM]-x-[DEA]-x(2)-[GA]-x-[GTA]-[SA]-x-G-H-x-[LIVM]-[GA]

SUBSTITUTE SHEET (RULE 26)

70

- [1] Golbeck J.H. Biochim. Biophys. Acta 895:167-204 (1987).  
 [2] Kjaerulff S., Andersen B., Nielsen V.S., Moller B.L.,  
 Okkels J.S. J. Biol. Chem. 268:18912-18916(1993).

5 17. Plant lipid transfer protein family signature (LTP)  
 Plant cells contain proteins, called lipid transfer  
 proteins (LTP) [1,2,3], which are able to facilitate the  
 transfer of phospholipids and other lipids across membranes.  
 These proteins, whose subcellular location is not yet known,  
 could play a major role in membrane biogenesis by conveying  
 phospholipids such as waxes or cutin from their site of  
 biosynthesis to membranes unable to form these lipids. Plant  
 LTP's are proteins of about 9 Kd (90 amino acids) which  
 contain eight conserved cysteine residues all involved in  
 disulfide bridges, as shown in the following schematic  
 representation.

```

+-----+ | +-----+ | | | |
*****
xCxxxxxCxxxxxCxxxxxCxxxxxCxxxxxCxxxxxCxxx | | | +-----
-----+ | +-----+

```

'C': conserved cysteine involved in a disulfide bond.  
 '\*': position of the pattern.

Consensus pattern: [LIVM]-[PA]-x(2)-C-x-[LIVM]-x-[LIVM]-x-  
 [LIVMFY]-x-[LIVM]-[ST]-x(3)-[DN]-C-x(2)-[LIVM] [The two C's  
 are involved in disulfide bonds]

- [1] Wirtz K.W.A. Annu. Rev. Biochem. 60:73-99(1991).  
 [2] Arondel V., Kader J.C. Experientia 46:579-585(1990).  
 [3] Ohlrogge J.B., Browse J., Somerville C.R. Biochim.  
 Biophys. Acta 1082:1-26(1991).

18. Ribosomal protein S7e signature

SUBSTITUTE SHEET (RULE 26)

71

A number of eukaryotic ribosomal proteins can be grouped  
 on the basis of sequence similarities [1]. One of these  
 families consists of:

- Mammalian S7.
- Xenopus S8.
- Insect S7.
- Yeast probable ribosomal protein S7 (N2212).
- Fission yeast probable ribosomal protein S7  
 (SpAC18G6.13c).

10 These proteins have about 200 amino acids. A highly conserved  
 stretch of 14 residues which is located in the central  
 section and which is rich in charged residues was selected as  
 a signature pattern.

15 Consensus pattern: [KR]-L-x-R-E-L-E-K-K-F-[SAP]-x-[KR]-H  
 [1] Salazar C.E., Mills-Hamm D.M., Kumar V., Collins F.H.  
 Nucleic Acids Res. 21:4147-4147(1993).

20 19. Ribosomal protein L34 signature  
 Ribosomal protein L34 is one of the proteins from the  
 large subunit of the prokaryotic ribosome. It is a small  
 basic protein of 44 to 51 amino-acid residues [1]. L34  
 belongs to a family of ribosomal proteins which, on the basis  
 of sequence similarities, groups: - Eubacterial L34.  
 - Red algal chloroplast L34. - Cyanelle L34.

A conserved region that corresponds to the N-terminal half of  
 L34 has been selected  
 as a signature pattern.

30 -Consensus pattern: K-[RG]-T-[FYWL]-[EQS]-x(5)-[KRHS]-x(4,5)-  
 G-F-x(2)-R  
 [ 1] Old I.G., Margarita D., Saint Girons I.  
 Nucleic Acids Res. 20:6097-6097(1992).

SUBSTITUTE SHEET (RULE 26)



72

## 20. Ribosomal protein L6 signatures

Ribosomal protein L6 is one of the proteins from the large ribosomal subunit. In *Escherichia coli*, L6 is known to bind directly to the 23S rRNA and is located at the aminoacyl-tRNA binding site of the peptidyltransferase center. It belongs to a family of ribosomal proteins which, on the basis of sequence similarities [1,2,3,4], groups: - Eubacterial L6.

- Algal chloroplast L6.
- Cyanelle L6.
- Archaeobacterial L6.
- Marchantia polymorpha mitochondrial L6.
- Yeast mitochondrial YmL6 (gene MREL6).
- Mammalian L9.
- Drosophila L9.
- Plants L9.
- Yeast L9 (Yl11).

While all the above proteins are evolutionary related it is very difficult to derive a pattern that will find them all. Two patterns were therefore created, the first to detect eubacterial, cyanelle and mitochondrial L6, the second to detect archaeobacterial L6 as well as eukaryotic L9.

-Consensus pattern: [PS]-[DENS]-x-y-k-[GA]-k-g-[LIVM]  
 -Consensus pattern: Q-x(3)-[LIVM]-x(2)-[KR]-x(2)-R-x-F-x-D-G-[LIVM]-Y-[LIVM]-x(2)-[KR]

[1] Suzuki K., Olivera J., Wool I.G. Gene 93:297-300(1990).

[2] Schwank S., Harrer R., Schueller H.-J., Schweizer E. Curr. Genet. 24:136-140(1993).

[3] Golden B.L., Ramakrishnan V., White S.W. EMBO J. 12:4901-4908(1993).

[4] Otake E., Hashimoto T., Mizuta K., Suzuki K. Protein Seq. Data Anal. 5:301-313(1993).

73

## 21. Ribosomal protein S14p/S29e (Ribosomal protein S14 signature)

Ribosomal protein S14 is one of the proteins from the small ribosomal subunit. In *Escherichia coli*, S14 is known to be required for the assembly of 30S particles and may also be responsible for determining the conformation of 16S rRNA at the A site. It belongs to a family of ribosomal proteins which, on the basis of sequence similarities [1,2], groups:

- Eubacterial S14.
- Algal and plant chloroplast S14.
- Cyanelle S14.
- Archaeobacterial Methanococcus vannielii S14.
- Plant mitochondrial S14.
- Yeast mitochondrial MRP2.
- Mammalian S29.
- Yeast YS29A/B.

S14 is a protein of 53 to 115 amino-acid residues. Our signature pattern is based on the few conserved positions located in the center of these proteins.

Consensus pattern: [RP]-x(0,1)-C-x(11,12)-[LIVMF]-x-[LIVMF]-[SC]-[RG]-x(3)-[RN]

[1] Chan Y.-L., Suzuki K., Olivera J., Wool I.G. Nucleic Acids Res. 21:649-655(1993).

[2] Otake E., Hashimoto T., Mizuta K. Protein Seq. Data Anal. 5:285-300(1993).

## 22. Ribosomal protein S16 signature

Ribosomal protein S16 is one of the proteins from the small ribosomal subunit. It belongs to a family of ribosomal proteins which, on the basis of sequence similarities [1], groups:

- Eubacterial S16.

- Algal and plant chloroplast S16.

- Cyanelle S16.

- Neurospora crassa mitochondrial S24 (cyt-21).

S16 is a protein of about 100 amino-acid residues. A

5 conserved region located in the N-terminal extremity of these proteins has been selected as a signature pattern.

Consensus pattern: [LIVMT]-x-[LIVM]-[KR]-L-[STAK]-R-x-G-[AKR]

10 [1] Otake E., Hashimoto T., Mizuta K. Protein Seq. Data Anal. 5:285-300(1993).

#### 23. Ribosomal protein S21 signature

15 Ribosomal protein S21 is one of the proteins from the small ribosomal subunit. So far S21 has only been found in eubacteria. It is a protein of 55 to 70 amino-acid residues. A conserved region in the N-terminal section of the protein has been selected as a signature pattern.

20 Consensus pattern: [DE]-x-A-[LIY]-[KR]-R-F-K-[KR]-x(3)-[KR]

#### 24. Universal stress protein family (Usp)

25 By a wide range of stress conditions members of the Usp family are predicted to be related to the MADS-box proteins transcript fact and bind to DNA [2]. Number of members: 39

[1] Expression and role of the universal stress protein, UspA, of *Escherichia coli* during growth arrest. Nystrom T, Neidhardt FC; Mol Microbiol 1994; 11:537-544.

30 [2] Sequence analysis of eukaryotic developmental proteins: ancient and novel domains. Mushegian AR, Koonin EV; Genetics 1996; 144:817-828.

### III. Methods of Modulating Polypeptide Production

5 Within the scope of invention are chimeric gene constructs wherein the promoter and the structural coding sequence and/or other regulatory sequences within said constructs are heterologous to each other. "Heterologous sequences" are those that are not operatively linked or are not contiguous to each other in nature. For example, a promoter from corn is considered heterologous to an *Arabidopsis* coding region sequence. Also, a promoter from a gene encoding a growth factor from corn is considered heterologous to a sequence encoding the corn receptor for the growth factor. Regulatory element sequences, such as UTRs or 3' end termination sequences that do not originate in nature from the same gene as the coding sequence originates from, are considered heterologous to said coding sequence. On the other hand, elements operatively linked in nature are not heterologous. Thus, the promoter and coding portion of a corn gene expressing an amino acid transporter are not heterologous to each other.

20 Such chimeric polynucleotides are of particular interest for modulating gene expression in a host cell upon transformation of said cell with said chimeric polynucleotide.

25 Also within the scope of the invention are DNA molecules, whereof at least a part or portion of these DNA molecules are presented in the Sequence Listing of the present application, and wherein the structural coding sequence is under the control of its own promoter and/or its own regulatory elements. Such DNA molecules are useful for transforming the genome of a host cell or an organism regenerated from said host cell.

30 Typically, such polynucleotides, whether chimeric or not, are "exogenous to" the genome of an individual host cell or the organism regenerated from said host cell, such as a plant cell, respectively for a plant, when initially or subsequently introduced into said host cell or organism, by any means other

than by a sexual cross. Examples of means by which this can be accomplished are described below, and include *Agrobacterium*-mediated transformation (of dicots - e.g. Salomon et al. *EMBO J.* 3:141 (1984); Herrera-Estrella et al. *EMBO J.* 2:987 (1983); A.C. Vergunst et al, *Nucleic Acids Res.* 26:11, 2729 (1998); of monocots, representative papers are those by Escudero et al., *Plant J.* 10:355 (1996), Ishida et al., *Nature Biotechnology* 14:745 (1996), May et al., *Bio/Technology* 13:486 (1995)), biolistic methods (Armaleo et al., *Current Genetics* 17:97 (1990)), electroporation, in planta techniques, and the like. Such a plant containing the exogenous nucleic acid is referred to here as an  $R_1$  generation transgenic plant. Transgenic plants which arise from a sexual cross with another parent line or by selfing are "descendants or the progeny" of a  $R_1$  plant and are generally called  $F_n$  plants or  $S_n$  plants, respectively,  $n$  meaning the number of generations.

The SDFs prepared as described herein can be used to prepare expression cassettes useful in a number of techniques for suppressing or enhancing expression.

### 20 III.A. Suppression

Expression cassettes of the invention can be used to suppress expression of endogenous genes which comprise the SDF sequence. Inhibiting expression can be useful, for instance, to tailor the ripening characteristics of a fruit (Oellier et al., *Science* 254:437 (1991)) or to influence seed size (W098/07842) or or to provoke cell ablation (Mariani et al., *Nature* 357: 384-387 (1992)).

As described above, a number of methods can be used to inhibit gene expression in plants, such as antisense, ribozyme, introduction of "exogenous" genes into a host cell, insertion of a polynucleotide sequence into the coding sequence and/or the promoter of the endogenous gene of interest, and the like.

### III.A.1. Antisense

An expression cassette as described above can be transformed into host cell or plant to produce an antisense strand of RNA. In plant cells, it has been suggested that antisense RNA inhibits gene expression by preventing the accumulation of mRNA which encodes the enzyme of interest, see, e.g., Sheehy et al., *Proc. Nat. Acad. Sci. USA*, 85:8805 (1988), and Hiatt et al., U.S. Patent No. 4,801,340.

### 10 III.A.2. Ribozymes

Similarly, ribozyme constructs can be transformed into a plant to cleave mRNA and down-regulate translation.

### 15 III.A.3. Co-Suppression

Another method of suppression is by introducing an exogenous copy of the gene to be suppressed. Introduction of expression cassettes in which a nucleic acid is configured in the sense orientation with respect to the promoter has been shown to be an effective means by which to block the transcription of target genes. A detailed description of this method is described above.

### 20 III.A.4. Insertion of Sequences into the Gene to be Modulated

Yet another means of suppressing gene expression is to insert a polynucleotide into the gene of interest to disrupt transcription or translation of the gene.

Homologous recombination could be used to target a polynucleotide insert to a gene using the Cre-Lox system (A.C. Vergunst et al., *Nucleic Acids Res.* 26:2729 (1998), A.C. Vergunst et al., *Plant Mol. Biol.* 38:393 (1998), H. Albert et al., *Plant J.* 7:649 (1995)).

78

In addition, random insertion of polynucleotides into a host cell genome can also be used to disrupt the gene of interest. Azpiroz-Leehan et al., *Trends in Genetics* 13:152 (1997). In this method, screening for clones from a library containing random insertions is preferred for identifying those that have polynucleotides inserted into the gene of interest. Such screening can be performed using probes and/or primers described above based on sequences from the Sequence Listing, fragments thereof, and substantially similar sequence thereto. The screening can also be performed by selecting clones or  $R_1$  plants having a desired phenotype.

### III.A.5. Promoter Modulation

Inactivation of the promoter that drives a gene of interest can modulate transcription and translation, and therefore expression. For example, triple helices can be formed using oligonucleotides based on sequences from the Sequence Listing, fragments thereof, and substantially similar sequence thereto. The oligonucleotide can be delivered to the host cell can bind to the promoter in the genome to form a triple helix and prevent transcription.

Additionally, a vector capable of producing the oligonucleotide can be inserted into the host cell to deliver the oligonucleotide.

### III.A.6. Expression of Mutants

An alternative method for inhibiting gene function is through the use of dominant negative mutations. Dominant negative mutations produce a mutant polypeptide which is capable of competing with the native polypeptide, but which does not produce the native result. Consequently, over expression of these mutations can titrate out an undesired activity of the native protein. For example, the inactive dominant-negative mutant may bind to the same receptor as the

79

native protein, preventing the native protein from activating a signal transduction pathway. Alternatively, the dominant-negative mutant can be an inactive enzyme still capable of binding to the same substrate as the native protein.

Dominant-negative mutants also can act upon the native protein itself to prevent activity. For example, the native protein may be active only as a homo-multimer or as one subunit of a hetero-multimer. Incorporation of an inactive subunit into the multimer with native subunit(s) can inhibit activity.

Thus, gene function can be modulated by insertion of an expression construct encoding a dominant-negative mutant into a host cell of interest.

### III.B. Enhanced Expression

Enhanced expression of a gene of interest in a host cell can be accomplished by either (1) insertion of an exogenous gene; or (2) promoter modulation.

#### III.B.1. Insertion of an Exogenous Gene

Insertion of an expression construct encoding an exogenous gene can boost the number of gene copies expressed in a host cell.

Such expression constructs can comprise genes that either encode the native protein that is of interest or that encode a variant that exhibits enhanced activity as compared to the native protein. Such genes encoding proteins of interest can be constructed from the sequences from the Sequence Listing, fragments thereof, and substantially similar sequence thereto.

Such an exogenous gene can include either a constitutive promoter permitting expression in any cell in a host organism or a promoter that directs expression only in

80

particular cells or times during a host cell life cycle or in response to environmental stimuli.

### III.B.2. Promoter Modulation

Some promoters require binding of a regulatory protein to be activated. Other promoters may need a protein that signals a promoter binding protein to expose a polymerase binding site. In either case, over-expression of such proteins can be used to enhance expression of a gene of interest by increasing the activation time of the promoter.

Such regulatory proteins are encoded by some of the sequences in the Sequence Listing, fragments thereof, and substantially similar sequences thereto.

Coding sequences for these proteins can be constructed as described above.

In some cases, duplication of enhancer elements or insertion of exogenous enhancer elements will increase expression of a desired gene from a particular promoter. The useful enhancer elements can be portions of one or more of the SDFs of the Sequence Listing.

15

### IV. Gene Constructs and Vector Construction

To use isolated SDFs of the present invention or a combination of them or parts and/or mutants and/or fusions of said SDFs in the above techniques, recombinant DNA vectors which comprise said SDFs and are suitable for transformation of cells, such as plant cells, are usually prepared.

The vector backbone can be any of those typical in the art such as plasmids, viruses, artificial chromosomes, BACs, YACs and PACs and vectors of the sort described by \*\*.

Typically, a vector will comprise the exogenous gene, which in its turn comprises an SDF of the present invention to be introduced into the genome of a host cell, and which gene may be an antisense construct, a ribozyme construct, or a

81

structural coding sequence with any desired transcriptional and/or translational regulatory sequences, such as promoters and 3' end termination sequences. Vectors of the invention can also include origins of replication, markers, homologous sequences, introns, etc.

A DNA sequence coding for the desired polypeptide, for example a cDNA sequence encoding a full length protein, will preferably be combined with transcriptional and translational initiation regulatory sequences which will direct the transcription of the sequence from the gene in the intended tissues of the transformed plant.

For example, for over-expression, a plant promoter fragment may be employed that will direct expression of the gene in all tissues of a regenerated plant. Such promoters are referred to herein as "constitutive" promoters and are active under most environmental conditions and states of development or cell differentiation. Examples of constitutive promoters include the cauliflower mosaic virus (CaMV) 35S transcription initiation region, the 1' or 2' promoter derived from T-DNA of *Agrobacterium tumefaciens*, and other transcription initiation regions from various plant genes known to those of skill.

Alternatively, the plant promoter may direct expression of an SDF of the invention in a specific tissue (tissue-specific promoters) or may be otherwise under more precise environmental control (inducible promoters). Examples of tissue-specific promoters under developmental control include promoters that initiate transcription only in certain tissues, such as root, ovule, fruit, seeds, or flowers. The promoter from a *LEC1* gene, described in copending application U.S. Ser. No. 09/103,478, is particularly useful for directing gene expression so that a desired gene product is located in embryos or seeds. Other suitable promoters include those from genes encoding storage proteins or the lipid body membrane protein, oleosin. A few root-specific promoters are noted

82

above. Examples of environmental conditions that may affect transcription by inducible promoters include anaerobic conditions, elevated temperature, or the presence of light.

If proper polypeptide expression is desired, a polyadenylation region at the 3'-end of the coding region should be included. The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from T-DNA.

The vector comprising the sequences (e.g., promoters or coding regions) from genes of the invention will typically comprise a marker gene that confers a selectable phenotype on plant cells. For example, the marker may encode biocide resistance, particularly antibiotic resistance, such as resistance to kanamycin, G418, bleomycin, hygromycin, or herbicide resistance, such as resistance to chlorosulfuron or phosphinotricin.

#### IV.A. Coding Sequences

Generally, the sequence in the transformation vector and to be introduced into the genome of the host cell does not need to be absolutely identical to an SDF of the present invention. Also, it is not necessary for it to be full length, relative to either the primary transcription product or fully processed mRNA. Use of sequences shorter than full-length may be preferred to avoid concurrent production of some plants that are overexpressors. Furthermore, the introduced sequence need not have the same intron or exon pattern as a native gene. Also, heterologous non-coding segments can be incorporated into the coding sequence without changing the desired amino acid sequence of the polypeptide to be produced.

#### 30 IV.B. Promoters

As explained above, introducing an exogenous SDF from the same species or an orthologous SDF from another species

83

can modulate the expression of a native gene corresponding to that SDF of interest. Such an SDF construct can be under the control of either a constitutive promoter (e.g., the promoter of the 35S gene of the cauliflower mosaic virus or the promoter of the gene encoding the cowpea trypsin inhibitor) or a highly regulated inducible promoter (e.g., a copper inducible promoter). The promoter of interest can initially be either endogenous or heterologous to the species in question. When re-introduced into the genome of said species, such promoter becomes "exogenous" to said species. The promoter-SDF construct can be made using standard recombinant DNA techniques (Sambrook et al. 1989) and can be introduced to the species of interest by *Agrobacterium*-mediated transformation or by other means of transformation (e.g., particle gun bombardment) as referenced above. Over-expression of an SDF transgene can lead to co-suppression of the homologous gene thereby creating some alterations in the phenotypes of the transformed species as demonstrated by similar analysis of the chalcone synthase gene (Napoli et al., *Plant Cell* 2:279 (1990) and van der Krol et al., *Plant Cell* 2:291 (1990)). If an SDF is found to encode a protein with desirable characteristics, its over-expression can be controlled so that its accumulation can be manipulated in an organ- or tissue-specific manner utilizing a promoter having such specificity.

Likewise, if the promoter of an SDF (or an SDF that includes a promoter) is found to be tissue-specific or developmentally regulated, such a promoter can be utilized to drive the expression of a specific gene of interest (e.g., seed storage protein or root-specific protein). Thus, the level of accumulation of a particular protein can be manipulated or its spatial localization in an organ- or tissue-specific manner can be altered.

#### IV. C Signal Peptides

In some cases it may be desirable for the protein encoded by an introduced exogenous or orthologous SDF to be targeted (1) to a particular organelle, (2) to interact with a particular molecule or (3) for secretion outside of the cell harboring the introduced SDF. This will be accomplished using a signal peptide.

Signal peptides direct protein targeting, are involved in ligand-receptor interactions and act in cell to cell communication. Many proteins, especially soluble proteins, contain a signal peptide that targets the protein to one of several different intracellular compartments. In plants, these compartments include, but are not limited to, the endoplasmic reticulum (ER), mitochondria, plastids (such as chloroplasts), the vacuole, the Golgi apparatus, protein storage vesicles (PSV) and, in general, membranes. Some signal peptide sequences are conserved, such as the Asn-Pro-Ile-Arg amino acid motif found in the N-terminal propeptide signal that targets proteins to the vacuole (Marty (1999) *The Plant Cell* 11: 587-599). Other signal peptides do not have a consensus sequence *per se*, but are largely composed of hydrophobic amino acids, such as those signal peptides targeting proteins to the ER (Vitale and Denecke (1999) *The Plant Cell* 11: 615-628). Still others do not appear to contain either a consensus sequence or an identified common secondary sequence, for instance the chloroplast stromal targeting signal peptides (Keegstra and Cline (1999) *The Plant Cell* 11: 557-570). Furthermore, some targeting peptides are bipartite, directing proteins first to an organelle and then to a membrane within the organelle (e.g. within the thylakoid lumen of the chloroplast; see Keegstra and Cline (1999) *The Plant Cell* 11: 557-570). In addition to the diversity in sequence and secondary structure, placement of the signal peptide is also varied. Proteins destined for

the vacuole, for example, have targeting signal peptides found at the N-terminus, at the C-terminus and at a surface location in mature, folded proteins.

Signal peptides also serve as ligands for some receptors. Perhaps the best known example of this is the interaction of the ER targeting signal peptide with the signal recognition particle (SRP). Here, the SRP binds to the signal peptide, halting translation, and the resulting SRP complex then binds to docking proteins located on the surface of the ER, prompting the transfer of the protein into the ER.

These characteristics of signal proteins can be used to more tightly control the expression of introduced SDFs. In particular, associating the appropriate signal sequence with a specific SDF can allow sequestering of the protein in specific organelles (plastids, as an example), secretion outside of the cell, targeting interaction with particular receptors, etc. Hence, the inclusion of signal proteins in constructs involving the SDFs of the invention increases the range of manipulation of SDF expression. To carry this out, constructs are made with the nucleotide sequence of a known signal peptide immediately 5' to the initiation of the coding region of an SDF so that the signal peptide is translated in frame with the coding region and immediately precedes it.

The nucleotide sequence of the signal peptide can be isolated from characterized genes using common molecular biological techniques or can be synthesized *in vitro*.

#### V. Transformation Techniques

A wide range of techniques for inserting exogenous polynucleotides are known for a number of host cells, including, without limitation, bacterial, yeast, mammalian, insect and plant cells.

86

Techniques for transforming a wide variety of higher plant species are well known and described in the technical and scientific literature. See, e.g. Weising et al., *Ann. Rev. Genet.* 22:421 (1988); and Cristou, *Euphytica*, v. 85, n.1-3:13-27, (1995).

DNA constructs of the invention may be introduced into the genome of the desired plant host by a variety of conventional techniques. For example, the DNA construct may be introduced directly into the genomic DNA of the plant cell using techniques such as electroporation and microinjection of plant cell protoplasts, or the DNA constructs can be introduced directly to plant tissue using ballistic methods, such as DNA particle bombardment. Alternatively, the DNA constructs may be combined with suitable T-DNA flanking regions and introduced into a conventional *Agrobacterium tumefaciens* host vector. The virulence functions of the *Agrobacterium tumefaciens* host will direct the insertion of the construct and adjacent marker into the plant cell DNA when the cell is infected by the bacteria (Vergunst et al., *Nucl. Acids. Res.* 26:2729 (1998) (site-directed integration using a Cre-Lox recombinase system); McCormac et al., *Mol. Biotechnol.* 8:199 (1997); Hamilton, *Gene* 200:107 (1997); Salomon et al. *EMBO J.* 3:141 (1984); Herrera-Estrella et al. *EMBO J.* 2:987 (1983).

Microinjection techniques are known in the art and well described in the scientific and patent literature. The introduction of DNA constructs using polyethylene glycol precipitation is described in Paszkowski et al. *EMBO J.* 3:2717 (1984). Electroporation techniques are described in Fromm et al. *Proc. Natl. Acad. Sci. USA* 82:5824 (1985). Ballistic transformation techniques are described in Klein et al. *Nature* 327:773 (1987). *Agrobacterium tumefaciens*-mediated transformation techniques, including disarming and use of binary vectors, are well described in the scientific literature. See, for example Hamilton, *CM., Gene* 200:107

87

(1997); Muller et al. *Mol. Gen. Genet.* 207:171 (1987); Komari et al. *Plant J.* 10:165 (1996); Venkateswarlu et al. *Biotechnology* 9:1103 (1991) and Gleave, AP, *Plant Mol. Biol.* 20:1203 (1992); Graves and Goldman, *Plant Mol. Biol.* 7:34 (1986) and Gould et al., *Plant Physiology* 95:426 (1991).

Transformed plant cells which are derived by any of the above transformation techniques can be cultured to regenerate a whole plant that possesses the transformed genotype and thus the desired phenotype such as seedlessness. Such regeneration techniques rely on manipulation of certain phytohormones in a tissue culture growth medium, typically relying on a biocide and/or herbicide marker which has been introduced together with the desired nucleotide sequences. Plant regeneration from cultured protoplasts is described in Evans et al., *Protoplasts Isolation and Culture* in "Handbook of Plant Cell Culture," pp. 124-176, MacMillan Publishing Company, New York, 1983; and Binding, *Regeneration of Plants, Plant Protoplasts*, pp. 21-73, CRC Press, Boca Raton, 1988. Regeneration can also be obtained from plant callus, explants, organs, or parts thereof. Such regeneration techniques are described generally in Klee et al. *Ann. Rev. of Plant Phys.* 38:467 (1987). Regeneration of monocots (rice) is described by Hosoyama et al. (*Biosci. Biotechnol. Biochem.* 58:1500 (1994)) and by Ghosh et al. (*J. Biotechnol.* 32:1 (1994)). The nucleic acids of the invention can be used to confer desired traits on essentially any plant.

Thus, the invention has use over a broad range of plants, including species from the genera *Asparagus*, *Atropa*, *Avena*, *Brassica*, *Citrus*, *Citrullus*, *Capsicum*, *Cucumis*, *Cucurbita*, *Daucus*, *Fragaria*, *Glycine*, *Gossypium*, *Helianthus*, *Heterocallis*, *Hordeum*, *Hyoscyamus*, *Lactuca*, *Linum*, *Lolium*, *Lycopersicon*, *Malus*, *Manihot*, *Majiorana*, *Medicago*, *Nicotiana*, *Oryza*, *Panicum*, *Panasetum*, *Persea*, *Pisum*, *Pyrus*, *Prunus*, *Raphanus*, *Secale*, *Senecio*, *Sinapis*, *Solanum*, *Sorghum*, *Trigonella*, *Triticum*, *Vitis*, *Vigna*, and, *Zea*.



88

One of skill will recognize that after the expression cassette is stably incorporated in transgenic plants and confirmed to be operable, it can be introduced into other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed.

The particular sequences of SDFs identified are provided in the attached the Sequence Listing. One of ordinary skill in the art, having this data, can obtain cloned DNA fragments, synthetic DNA fragments or polypeptides constituting desired sequences by recombinant methodology known in the art or described herein.

#### EXAMPLES

The invention is illustrated by way of the following examples. The invention is not limited by these examples as the scope of the invention is defined solely by the claims following.

#### EXAMPLE 1: SOUTHERN HYBRIDIZATIONS

The SDFs of the invention can be used in Southern hybridizations as described above. The following describes extraction of DNA from nuclei of plant cells, digestion of the nuclear DNA and separation by length, transfer of the separated fragments to membranes, preparation of probes for hybridization, hybridization and detection of the hybridized probe.

The procedures described herein can be used to isolate related polynucleotides or for diagnostic purposes. Moderate stringency hybridization conditions, as defined above, are described in the present example. These conditions result in detection of hybridization between sequences having at least 70% sequence identity. As described above, the hybridization and wash conditions can be changed to reflect the desired

89

degree of sequence identity between probe and target sequences that can be detected.

In the following procedure, a probe for the hybridization is produced from two PCR reactions using two primers from genomic sequence of *Arabidopsis thaliana*. As described above, the particular template for generating the probe can be any desired template.

The first PCR product is assessed to validate the size of the primer to assure it is of the expected size. Then the product of the first PCR is used as a template, with the same pair of primers used in the first PCR, in a second PCR that produces a labeled product used as the probe.

Fragments detected by hybridization, or other bands of interest, can be isolated from gels used to separate genomic DNA fragments by known methods for further purification and/or characterization.

#### Buffers for nuclear DNA extraction

##### 1. 10X HB

	1000 ml	
40 mM spermidine	10.2 g	Spermine (Sigma S-2876) and spermidine (Sigma S-2501)
10 mM spermine	3.5 g	Stabilize chromatin and the nuclear membrane
0.1 M EDTA (disodium)	37.2 g	EDTA inhibits nuclease
0.1 M Tris	12.1 g	Buffer
0.8 M KCl	59.6 g	Adjusts ionic strength for stability of nuclei

Adjust pH to 9.5 with 10 N NaOH. It appears that there is a nuclease present in leaves. Use of pH 9.5 appears to inactivate this nuclease.

90

2. 2 M sucrose (684 g per 1000 ml)  
Heat about half the final volume of water to about 50°C.  
Add the sucrose slowly then bring the mixture to close  
to final volume; stir constantly until it has dissolved.  
Bring the solution to volume.

5

3. Sarkosyl solution (lyses nuclear membranes)

	<u>1000 ml</u>	
N-lauroyl sarcosine (Sarkosyl)	20.0 g	
0.1 M Tris	12.1 g	
0.04 M EDTA (Disodium)	14.9 g	

10

Adjust the pH to 9.5 after all the components are  
dissolved and bring up to the proper volume.

4. 20% Triton X-100  
80 ml Triton X-100

15

320 ml 1xHB (w/o  $\beta$ -ME and PMSF)

Prepare in advance; Triton takes some time to dissolve

#### A. Procedure

1. Prepare 1X "H" buffer (keep ice-cold during use)

	<u>1000 ml</u>	
10X HB	100 ml	
2 M sucrose	250 ml a non-ionic osmoticum	
Water	634 ml	

20

#### Added just before use:

100 mM PMSF\*  
10 ml a protease  
inhibitor; protects  
nuclear membrane proteins

25

91

$\beta$ -mercaptoethanol 1 ml inactivates nuclease  
by reducing disulfide  
bonds

\*100 mM PMSF

- 5 (phenyl methyl sulfonyl fluoride, Sigma P-7626)  
(add 0.0875 g to 5 ml 100% ethanol)

2. Homogenize the tissue in a blender (use 300-400 ml of  
1xHB per blender). Be sure that you use 5-10 ml of HB  
buffer per gram of tissue. Blenders generate heat so  
be sure to keep the homogenate cold. It is necessary to  
put the blenders in ice periodically.

10

3. Add the 20% Triton X-100 (25 ml per liter of homogenate)  
and gently stir on ice for 20 min. This lyses plastid,  
but not nuclear, membranes.

15

4. Filter the tissue suspension through several nylon  
filters into an ice-cold beaker. The first filtration  
is through a 250-micron membrane; the second is through  
an 85-micron membrane; the third is through a 50-micron  
membrane; and the fourth is through a 20-micron  
membrane. Use a large funnel to hold the filters.  
Filtration can be sped up by gently squeezing the liquid  
through the filters.

20

5. Centrifuge the filtrate at 1200 x g for 20 min. at 4°C to  
pellet the nuclei.

- 25 6. Discard the dark green supernatant. The pellet will  
have several layers to it. One is starch; it is white  
and gritty. The nuclei are gray and soft. In the early

92

steps, there may be a dark green and somewhat viscous layer of chloroplasts.

Wash the pellets in about 25 ml cold H buffer (with Triton X-100) and resuspend by swirling gently and pipetting. After the pellets are resuspended.

Pellet the nuclei again at 1200 - 1300 x g. Discard the supernatant.

Repeat the wash 3-4 times until the supernatant has changed from a dark green to a pale green. This usually happens after 3 or 4 resuspensions. At this point, the pellet should be grayish white and very slippery. The Triton X-100 in these repeated steps helps to destroy the chloroplasts and mitochondria that contaminate the prep.

Resuspend the nuclei for a final time in a total of 15 ml of H buffer and transfer the suspension to a sterile 125 ml Erlenmeyer flask.

7. Add 15 ml, dropwise, cold 2% Sarkosyl, 0.1 M Tris, 0.04 M EDTA solution (pH 9.5) while swirling gently. This lyses the nuclei. The solution will become very viscous.

8. Add 30 grams of CsCl and gently swirl at room temperature until the CsCl is in solution. The mixture will be gray, white and viscous.

9. Centrifuge the solution at 11,400 x g at 4°C for at least 30 min. The longer this spin is, the firmer the protein pellicle.

93

10. The result should be a clear green supernatant over a white pellet, and (perhaps) under a protein pellicle. Carefully remove the solution under the protein pellicle and above the pellet. Determine the density of the solution by weighing 1 ml of solution and add CsCl if necessary to bring to 1.57 g/ml. The solution contains dissolved solids (sucrose etc) and the refractive index alone will not be an accurate guide to CsCl concentration.

11. Add 20 µl of 10 mg/ml EtBr per ml of solution.

12. Centrifuge at 184,000 x g for 16 to 20 hours in a fixed-angle rotor.

13. Remove the dark red supernatant that is at the top of the tube with a plastic transfer pipette and discard. Carefully remove the DNA band with another transfer pipette. The DNA band should be visible in room light; otherwise, use a long wave UV light to locate the band.

14. Extract the ethidium bromide with isopropanol saturated with water and salt. Once the solution is clear, extract at least two more times to ensure that all of the EtBr is gone. Be very gentle, as it is very easy to shear the DNA at this step. This extraction may take a while because the DNA solution tends to be very viscous. If the solution is too viscous, dilute it with TE.

15. Dialyze the DNA for at least two days against several changes (at least three times) of TE (10 mM Tris, 1mM EDTA, pH 8) to remove the cesium chloride.

94

16. Remove the dialyzed DNA from the tubing. If the dialyzed DNA solution contains a lot of debris, centrifuge the DNA solution at least at 2500 x g for 10 min. and carefully transfer the clear supernatant to a new tube. Read the A260 concentration of the DNA.

5

17. Assess the quality of the DNA by agarose gel electrophoresis (1% agarose gel) of the DNA. Load 50 ng and 100 ng (based on the OD reading) and compare it with known and good quality DNA. Undigested lambda DNA and a lambda-HindIII-digested DNA are good molecular weight makers.

10

#### Protocol for Digestion of Genomic DNA

##### Protocol:

1. The relative amounts of DNA for different crop plants that provide approximately a balanced number of genome equivalent is given in Table 3. Note that due to the size of the wheat genome, wheat DNA will be underrepresented. Lambda DNA provides a useful control for complete digestion.

15

2. Precipitate the DNA by adding 3 volumes of 100% ethanol. Incubate at -20°C for at least two hours. Yeast DNA can be purchased and made up at the necessary concentration, therefore no precipitation is necessary for yeast DNA.

20

3. Centrifuge the solution at 11,400 x g for 20 min. Decant the ethanol carefully (be careful not to disturb the pellet). Be sure that the residual ethanol is completely removed either by vacuum desiccation or by carefully wiping the sides of the tubes with a clean tissue.

25

95

4. Resuspend the pellet in an appropriate volume of water. Be sure the pellet is fully resuspended before proceeding to the next step. This may take about 30 min.

5

5. Add the appropriate volume of 10X reaction buffer provided by the manufacturer of the restriction-enzyme to the resuspended DNA followed by the appropriate volume of enzymes. Be sure to mix it properly by slowly swirling the tubes.

10

6. Set-up the lambda digestion-control for each DNA that you are digesting.

7.

Incubate both the experimental and lambda digests overnight at 37°C. Spin down condensation in a microfuge before proceeding.

15

8. After digestion, add 2 µl of loading dye (typically 0.25% bromophenol blue, 0.25% xylene cyanol in 15% Ficoll or 30% glycerol) to the lambda-control digests and load in 1% TPE-agarose gel (TPE is 90 mM Tris-phosphate, 2 mM EDTA, pH 8). If the lambda DNA in the lambda control digests are completely digested, proceed with the precipitation of the genomic DNA in the digests.

20

9. Precipitate the digested DNA by adding 3 volumes of 100% ethanol and incubating in -20°C for at least 2 hours (preferably overnight).

25

EXCEPTION: *Arabidopsis* and yeast DNA are digested in an appropriate volume; they don't have to be precipitated.

10. Resuspend the DNA in an appropriate volume of TE (e.g., 22 µl x 50 blots = 1100 µl) and an appropriate volume of 10X loading dye (e.g., 2.4 µl x 50 blots = 120 µl). Be careful in pipetting the loading dye - it is viscous. Be sure you are pipetting the correct volume.

5

Table 3

Some guide points in digesting genomic DNA.

Species	Genome Size	Size Relative to Arabidopsis	Genome Equivalent to 2 µg Arabidopsis DNA	Amount of DNA per blot
Arabidopsis	120 Mb	1X	1X	2 µg
Brassica	1,100 Mb	9.2X	0.54X	10 µg
Corn	2,800 Mb	23.3X	0.43X	20 µg
Cotton	2,300 Mb	19.2X	0.52X	20 µg
Oat	11,300 Mb	94X	0.11X	20 µg
Rice	400 Mb	3.3X	0.75X	5 µg
Soybean	1,100 Mb	9.2X	0.54X	10 µg
Sugarbeet	758 Mb	6.3X	0.8X	10 µg
Sweetclover	1,100 Mb	9.2X	0.54X	10 µg
Wheat	16,000 Mb	133X	0.08X	20 µg
Yeast	15 Mb	0.12X	1X	0.25 µg

#### 10 Protocol for Southern Blot Analysis

The digested DNA samples are electrophoresed in 1% agarose gels in 1x TBE buffer. Low voltage; overnight separations are preferred. The gels are stained with EtBr and photographed.

15

1. For blotting the gels, first incubate the gel in 0.25 N HCl (with gentle shaking) for about 15 min.

2. Then briefly rinse with water. The DNA is denatured by 2 incubations. Incubate (with shaking) in 0.5 M NaOH in 1.5 M NaCl for 15 min.

3. The gel is then briefly rinsed in water and neutralized by incubating twice (with shaking) in 1.5 M Tris pH 7.5 in 1.5 M NaCl for 15 min.

5

4. A nylon membrane is prepared by soaking it in water for at least 5 min, then in 6X SSC for at least 15 min. before use. (20x SSC is 175.3 g NaCl, 88.2 g sodium citrate per liter, adjusted to pH 7.0.)

10

5. The nylon membrane is placed on top of the gel and all bubbles in between are removed. The DNA is blotted from the gel to the membrane using an absorbent medium, such as paper toweling and 6x SSC buffer. After the transfer, the membrane may be lightly brushed with a gloved hand to remove any agarose sticking to the surface.

15

6. The DNA is then fixed to the membrane by UV crosslinking and baking at 80°C. The membrane is stored at 4°C until use.

20

#### B. Protocol for PCR Amplification of Genomic Fragments in Arabidopsis

##### Amplification procedures:

1. Mix the following in a 0.20 ml PCR tube or 96-well PCR plate:

25

Volume	Stock	Final Amount or Conc.
--------	-------	-----------------------

98

0.5 µl	~ 10 ng/µl genomic DNA <sup>1</sup>	5 ng
2.5 µl	10X PCR buffer	20 mM Tris, 50 mM KCl
0.75 µl	50 mM MgCl <sub>2</sub>	1.5 mM
1 µl	10 pmol/µl Primer 1 (Forward)	10 pmol
1 µl	10 pmol/µl Primer 2 (Reverse)	10 pmol
0.5 µl	5 mM dNTPs	0.1 mM
0.1 µl	5 units/µl Platinum Taq™ (Life Technologies, Gaithersburg, MD) DNA Polymerase	1 units
(to 25 µl)	Water	

2. The template DNA is amplified using a Perkin Elmer 9700 PCR machine:

- 1) 94°C for 10 min. followed by

2)	3)	4)
5 cycles:	5 cycles:	25 cycles:
94 °C - 30 sec	94 °C - 30 sec	94 °C - 30 sec
62 °C - 30 sec	58 °C - 30 sec	53 °C - 30 sec
72 °C - 3 min	72 °C - 3 min	72 °C - 3 min

- 5) 72°C for 7 min. Then the reactions are stopped by chilling to 4°C.

<sup>1</sup> Arabidopsis DNA is used in the present experiment, but the procedure is a general one.

99

The procedure can be adapted to a multi-well format if necessary.

#### Quantification and Dilution of PCR Products:

1. The product of the PCR is analyzed by electrophoresis in a 1% agarose gel. A linearized plasmid DNA can be used as a quantification standard (usually at 50, 100, 200, and 400 ng). These will be used as references to approximate the amount of PCR products. HindIII-digested Lambda DNA is useful as a molecular weight marker. The gel can be run fairly quickly; e.g., at 100 volts. The standard gel is examined to determine that the size of the PCR products is consistent with the expected size and if there are significant extra bands or smeary products in the PCR reactions.

2. The amounts of PCR products can be estimated on the basis of the plasmid standard.

3. For the small number of reactions that produce extraneous bands, a small amount of DNA from bands with the correct size can be isolated by dipping a sterile 10-µl tip into the band while viewing through a UV Transilluminator. The small amount of agarose gel (with the DNA fragment) is used in the labeling reaction.

#### C. Protocol for PCR-DIG-Labeling of DNA

##### Solutions:

- 25 Reagents in PCR reactions (diluted PCR products, 10X PCR Buffer, 50 mM MgCl<sub>2</sub>, 5 U/µl Platinum Taq Polymerase, and the primers)

10X dNTP + DIG-11-dUTP [1:5]: (2 mM dATP, 2 mM dCTP, 2 mM dGTP, 1.65 mM dTTP, 0.35 mM DIG-11-dUTP)

100

10X dNTP + DIG-11-dUTP [1:10]: (2 mM dATP, 2 mM dCTP, 2 mM dGTP, 1.81 mM dTTP, 0.19 mM DIG-11-dUTP)

10X dNTP + DIG-11-dUTP [1:15]: (2 mM dATP, 2 mM dCTP, 2 mM dGTP, 1.875 mM dTTP, 0.125 mM DIG-11-dUTP)

5 TE buffer (10 mM Tris, 1 mM EDTA, pH 8)

Maleate buffer: In 700 ml of deionized distilled water, dissolve 11.61 g maleic acid and 8.77 g NaCl. Add NaOH to adjust the pH to 7.5. Bring the volume to 1 L. Stir for 15 min. and sterilize.

10 10% blocking solution: In 80 ml deionized distilled water, dissolve 1.16g maleic acid. Next, add NaOH to adjust the pH to 7.5. Add 10 g of the blocking reagent powder (Boehringer Mannheim, Indianapolis, IN, Cat. no. 1096176). Heat to 60°C while stirring to dissolve the powder. Adjust the volume to 100 ml with water. Stir and sterilize.

1% blocking solution: Dilute the 10% stock to 1% using the maleate buffer.

20 Buffer 3 (100 mM Tris, 100 mM NaCl, 50 mM MgCl<sub>2</sub>, pH9.5). Prepared from autoclaved solutions of 1M Tris pH 9.5, 5 M NaCl, and 1 M MgCl<sub>2</sub> in autoclaved distilled water.

101

# Procedure:

1. PCR reactions are performed in 25 µl volumes containing:

PCR buffer 1X  
MgCl<sub>2</sub> 1.5 mM

5 10X dNTP + DIG-11-dUTP 1X (please see the note below)

Platinum Taq™ Polymerase 1 unit  
10 pg probe DNA  
10 pmol primer 1

10 **Note:**

Use for:

10X dNTP + DIG-11-dUTP (1:5)	< 1 kb
10X dNTP + DIG-11-dUTP (1:10)	1 kb to 1.8 kb
10X dNTP + DIG-11-dUTP (1:15)	> 1.8 kb

2. The PCR reaction uses the following amplification cycles:

1) 94°C for 10 min.

2)	3)	4)
5 cycles:	5 cycles:	25 cycles:
95°C - 30 sec	95°C - 30 sec	95°C - 30 sec
61°C - 1 min	59°C - 1 min	51°C - 1 min
73°C - 5 min	75°C - 5 min	73°C - 5 min

5) 72°C for 8 min. The reactions are terminated by chilling to 4°C (hold).

3. The products are analyzed by electrophoresis- in a 1% agarose gel, comparing to an aliquot of the unlabelled probe starting material.

102

4. The amount of DIG-labeled probe is determined as follows:

Make serial dilutions of the diluted control DNA in dilution buffer (TE: 10 mM Tris and 1 mM EDTA, pH 8) as shown in the following table:

DIG-labeled control DNA starting conc.	Stepwise Dilution	Final Conc. (Dilution Name)
5 ng/ $\mu$ l	1 $\mu$ l in 49 $\mu$ l TE	100 pg/ $\mu$ l (A)
100 pg/ $\mu$ l (A)	25 $\mu$ l in 25 $\mu$ l TE	50 pg/ $\mu$ l (B)
50 pg/ $\mu$ l (B)	25 $\mu$ l in 25 $\mu$ l TE	25 pg/ $\mu$ l (C)
25 pg/ $\mu$ l (C)	20 $\mu$ l in 30 $\mu$ l TE	10 pg/ $\mu$ l (D)

a. Serial deletions of a DIG-labeled standard DNA ranging from 100 pg to 10 pg are spotted onto a positively charged nylon membrane, marking the membrane lightly with a pencil to identify each dilution.

b. Serial dilutions (e.g., 1:50, 1:2500, 1:10,000) of the newly labeled DNA probe are spotted.

c. The membrane is fixed by UV crosslinking.

d. The membrane is wetted with a small amount of maleate buffer and then incubated in 1% blocking solution for 15 min at room temp.

e. The labeled DNA is then detected using alkaline phosphatase conjugated anti-DIG antibody (Boehringer Mannheim, Indianapolis, IN, cat. no.

103

1093274) and an NBT substrate according to the manufacture's instruction.

f. Spot intensities of the control and experimental dilutions are then compared to estimate the concentration of the PCR-DIG-labeled probe.

#### D. Prehybridization and Hybridization of Southern Blots

##### Solutions:

100% Formamide purchased from Gibco

20X SSC (1X = 0.15 M NaCl, 0.015 M Na<sub>3</sub>citrate)  
per L: 175 g NaCl  
87.5 g Na<sub>3</sub>citrate·2H<sub>2</sub>O

20% Sarkosyl (N-lauroyl-sarcosine)

20% SDS (sodium dodecyl sulphate)

10% Blocking Reagent: In 80 ml deionized distilled water, dissolve 1.16 g maleic acid. Next, add NaOH to adjust the pH to 7.5. Add 10 g of the blocking reagent powder. Heat to 60°C while stirring to dissolve the powder. Adjust the volume to 100 ml with water. Stir and sterilize.

##### Prehybridization Mix:

Final Concentration	Components	Volume (per 100 ml)	Stock
50%	Formamide	50 ml	100%
5X	SSC	25 ml	20X
0.1%	Sarkosyl	0.5 ml	20%



104

0.02%	SDS	0.1 ml	20%
2%	Blocking Reagent	20 ml	10%
	Water	4.4 ml	

General Procedures:

1. Place the blot in a heat-sealable plastic bag and add an appropriate volume of prehybridization solution (30 ml/100cm<sup>2</sup>) at room temperature. Seal the bag with a heat sealer, avoiding bubbles as much as possible. Lay down the bags in a large plastic tray (one tray can accommodate at least 4-5 bags). Ensure that the bags are lying flat in the tray so that the prehybridization solution is evenly distributed throughout the bag. Incubate the blot for at least 2 hours with gentle agitation using a waver shaker.

10

2. Denature DIG-labeled DNA probe by incubating for 10 min. at 98°C using the PCR machine and immediately cool it to 4°C.

3. Add probe to prehybridization solution (25 ng/ml; 30 ml = 750 ng total probe) and mix well but avoid foaming. Bubbles may lead to background.

15

4. Pour off the prehybridization solution from the hybridization bags and add new prehybridization and probe solution mixture to the bags containing the membrane.

20

5. Incubate with gentle agitation for at least 16 hours.

6. Proceed to medium stringency post-hybridization wash:

105

Three times for 20 min. each with gentle agitation using 1X SSC, 1% SDS at 60°C.

All wash solutions must be prewarmed to 60°C. Use about 100 ml of wash solution per membrane.

- 5 To avoid background keep the membranes fully submerged to avoid drying in spots; agitate sufficiently to avoid having membranes stick to one another.

7. After the wash, proceed to immunological detection and CSPD development.

**10 E. Procedure for Immunological Detection with CSPD**

Solutions:

Buffer 1:

Maleic acid buffer (0.1 M maleic acid, 0.15 M NaCl; adjusted to pH 7.5 with NaOH)

- 15 Washing buffer:

Maleic acid buffer with 0.3% (v/v) Tween 20.

Blocking stock solution 10% blocking reagent in buffer 1.

Dissolve (10X concentration):  
blocking reagent powder (Boehringer Mannheim, Indianapolis, IN, cat. no. 1096176) by constantly stirring on a 65°C heating block or heat in a microwave, autoclave and store at 4°C.

- 25 Buffer 2

106

(1X blocking solution): Dilute the stock solution 1:10 in Buffer 1.

Detection buffer: 0.1 M Tris, 0.1 M NaCl, pH 9.5

Procedure:

5 1. After the post-hybridization wash the blots are briefly rinsed (1-5 min.) in the maleate washing buffer with gentle shaking.

2. Then the membranes are incubated for 30 min. in Buffer 2 with gentle shaking.

10 3. Anti-DIG-AP conjugate (Boehringer Mannheim, Indianapolis, IN, cat. no. 1093274) at 75 mU/ml (1:10,000) in Buffer 2 is used for detection. 75 ml of solution can be used for 3 blots.

15 4. The membrane is incubated for 30 min. in the antibody solution with gentle shaking.

5. The membrane are washed twice in washing buffer with gentle shaking. About 250 mls is used per wash for 3 blots.

20 6. The blots are equilibrated for 2-5 min in 60 ml detection buffer.

7. Dilute CSPD (1:200) in detection buffer. (This can be prepared ahead of time and stored in the dark at 4°C).

25 The following steps must be done individually. Bags (one for detection and one for exposure) should be cut and ready before doing the following steps.

107

8. The blot is carefully removed from the detection buffer and excess liquid removed without drying the membrane. The blot is immediately placed in a bag and 1.5 ml of CSPD solution is added. The CSPD solution can be spread over the membrane. Bubbles present at the edge and on the surface of the blot should be removed by gentle rubbing. The membrane is incubated for 5 min. in CSPD solution.

5 9. Excess liquid is removed and the membrane is blotted briefly (DNA side up) on Whatman 3MM paper. Do not let the membrane dry completely.

10 10. Seal the damp membrane in a hybridization bag and incubate for 10 min at 37°C to enhance the luminescent reaction.

15 11. Expose for 2 hours at room temperature to X-ray film. Multiple exposures can be taken. Luminescence continues for at least 24 hours and signal intensity increases during the first hours.

**Example 2: Transformation of Carrot Cells**

20 Transformation of plant cells can be accomplished by a number of methods, as described above. Similarly, a number of plant genera can be regenerated from tissue culture following transformation. Transformation and regeneration of carrot cells as described herein is illustrative.

25 Single cell suspension cultures of carrot (*Daucus carota*) cells are established from hypocotyls of cultivar Early Nantes in B<sub>5</sub> growth medium (O.L. Gamborg et al., *Plant Physiol.* 45:372 (1970)) plus 2,4-D and 15 mM CaCl<sub>2</sub> (B<sub>5</sub> -44 medium) by methods known in the art. The suspension cultures are subcultured by adding 10 ml of the suspension culture to

108

40 ml of B<sub>3</sub>-44 medium in 250 ml flasks every 7 days and are maintained in a shaker at 150 rpm at 27 °C in the dark.

The suspension culture cells are transformed with exogenous DNA as described by Z. Chen et al. *Plant Mol. Bio.* 36:163 (1998). Briefly, 4-days post-subculture cells are incubated with cell wall digestion solution containing 0.4 M sorbitol, 2% driselase, 5mM MES (2-[N-Morpholino] ethanesulfonic acid) pH 5.0 for 5 hours. The digested cells are pelleted gently at 60 xg for 5 min. and washed twice in W5 solution containing 154 mM NaCl, 5 mM KCl, 125 mM CaCl<sub>2</sub> and 5mM glucose, pH 6.0. The protoplasts are suspended in MC solution containing 5 mM MES, 20 mM CaCl<sub>2</sub>, 0.5 M mannitol, pH 5.7 and the protoplast density is adjusted to about  $4 \times 10^6$  protoplasts per ml.

15 15-60 µg of plasmid DNA is mixed with 0.9 ml of protoplasts. The resulting suspension is mixed with 40% polyethylene glycol (MW 8000, PEG 8000), by gentle inversion a few times at room temperature for 5 to 25 min. Protoplast culture medium known in the art is added into the PEG-DNA-protoplast mixture. Protoplasts are incubated in the culture medium for 24 hour to 5 days and cell extracts can be used for assay of transient expression of the introduced gene. Alternatively, transformed cells can be used to produce transgenic callus, which in turn can be used to produce transgenic plants, by methods known in the art. See, for example, Nomura and Komamine, *Plt. Phys.* 79:988-991 (1985), *Identification and Isolation of Single Cells that Produce Somatic Embryos in Carrot Suspension Cultures.*

The invention being thus described, it will be apparent to one of ordinary skill in the art that various modifications of the materials and methods for practicing the invention can be made. Such modifications are to be

109

considered within the scope of the invention as defined by the following claims.

Each of the references from the patent and periodical literature cited herein is hereby expressly incorporated in its entirety by such citation.

110

**TABLE 1****Maximum Length Sequence:**

related to:

Clone IDs:

5

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 1

- Ceres seq\_id 1007546

- Alternative transcription start site(s) located in SEQ

10

ID NO 1:

-96,-51,25,27,29,55,64

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 2

- Ceres seq\_id 1007547

- Location of start within SEQ ID NO 1: at 1 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 2: at 32 aa.

20

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 1

- gi No. 3879939

- Description:

- % Identity: 45.9

- Alignment Length: 61

- Location of Alignment in SEQ ID NO 2: from 48 to

108

30

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 3

- Ceres seq\_id 1007548

- Location of start within SEQ ID NO 1: at 100 nt.

35

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 2

- gi No. 3879939

- Description:

- % Identity: 45.9

- Alignment Length: 61

- Location of Alignment in SEQ ID NO 3: from 15 to

45

75

40

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 4

- Ceres seq\_id 1007549

- Location of start within SEQ ID NO 1: at 121 nt.

50

111

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 3

- gi No. 3879939

- Description:

- % Identity: 45.9

- Alignment Length: 61

- Location of Alignment in SEQ ID NO 4: from 8 to 68

10

(Ba) Polypeptide Activities: Similar to yeast membrane protein activities

15

Maximum Length Sequence:

related to:

Clone IDs:

9568

402131

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 5

- Ceres seq\_id 1007583

- Alternative transcription start site(s) located in SEQ

ID NO 5:

2,3,4,7,10,11,12,17,42,43,300,505

25

20

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 6

- Ceres seq\_id 1007584

- Location of start within SEQ ID NO 5: at 55 nt.

30

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- KH domain

- Location within SEQ ID NO 6: from 47 to 95 aa.

35

40

(Dp) Related Amino Acid Sequences

- Alignment No. 4

- gi No. 133940

- Description:

- % Identity: 75.1

- Alignment Length: 250

- Location of Alignment in SEQ ID NO 6: from 1 to

249

45

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 7

- Ceres seq\_id 1007585

- Location of start within SEQ ID NO 5: at 184 nt.

50

112

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- KH domain
- Location within SEQ ID NO 7: from 4 to 52 aa.

5

## (Dp) Related Amino Acid Sequences

- Alignment No. 5
- gi No. 133940
- Description:
- % Identity: 75.1
- Alignment Length: 250
- Location of Alignment in SEQ ID NO 7: from 1 to

206

## Maximum Length Sequence:

related to:

Clone IDs:

91769

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 8
- Ceres seq\_id 1008148
- Alternative transcription start site(s) located in SEQ

20

ID NO 8:

-19,2,3,4,5,6,7,9,10,11,12,14

25

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 9
- Ceres seq\_id 1008149
- Location of start within SEQ ID NO 8: at 3 nt.

30

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 6
- gi No. 4539292
- Description:
- % Identity: 89.9
- Alignment Length: 181
- Location of Alignment in SEQ ID NO 9: from 25 to

40

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 10
- Ceres seq\_id 1008150
- Location of start within SEQ ID NO 8: at 75 nt.

45

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 7
- gi No. 4539292

50

113

- Description:
- % Identity: 89.9
- Alignment Length: 181
- Location of Alignment in SEQ ID NO 10: from 1 to

5

179

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 11
- Ceres seq\_id 1008151
- Location of start within SEQ ID NO 8: at 210 nt.

10

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 8
- gi No. 4539292
- Description:
- % Identity: 89.9
- Alignment Length: 181
- Location of Alignment in SEQ ID NO 11: from 1 to

20

134

(Ba) Polypeptide Activities: Similar to 40S Ribosomal protein activities, and glycine rich RNA binding protein activities.

25

## Maximum Length Sequence:

related to:

Clone IDs:

8898

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 12
- Ceres seq\_id 1008334
- Alternative transcription start site(s) located in SEQ

35

ID NO 12:

-12,29,30

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 13
- Ceres seq\_id 1008335
- Location of start within SEQ ID NO 12: at 2 nt.

40

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 9
- gi No. 3602948
- Description:
- % Identity: 43.7
- Alignment Length: 71

45

50

114

- Location of Alignment in SEQ ID NO 13: from 178 to 248

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 14
- Ceres seq\_id 1008336
- Location of start within SEQ ID NO 12: at 59 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 14: at 22 aa.

10

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 10
- gi No. 3602948
- Description:
- % Identity: 43.7
- Alignment Length: 71
- Location of Alignment in SEQ ID NO 14: from 159 to 229

20

(Ba) Polypeptide Activities: Plant specific gene, Chloroplast specific gene.

25

## Maximum Length Sequence:

related to:

Clone IDs:

8286

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 15
- Ceres seq\_id 1008701
- Alternative transcription start site(s) located in SEQ ID NO 15: -6,2,3,5,6,7,18,24,25,28,31,33,35,37,42,50

35

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 16
- Ceres seq\_id 1008702
- Location of start within SEQ ID NO 15: at 1 nt.

40

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 11
- gi No. 1083282
- Description:
- % Identity: 41.4
- Alignment Length: 106
- Location of Alignment in SEQ ID NO 16: from 23 to 121

50

115

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 17
- Ceres seq\_id 1008703
- Location of start within SEQ ID NO 15: at 67 nt.

5

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 12
- gi No. 1083282
- Description:
- % Identity: 41.4
- Alignment Length: 106
- Location of Alignment in SEQ ID NO 17: from 1 to 99

15

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 18
- Ceres seq\_id 1008704
- Location of start within SEQ ID NO 15: at 2 nt.

20

(Ba) Polypeptide Activities: Similar to cytochrome C oxidase activities.

25

## Maximum Length Sequence:

related to:

Clone IDs:

7792

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 19
- Ceres seq\_id 1009003
- Alternative transcription start site(s) located in SEQ ID NO 19: 2,374

35

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 20
- Ceres seq\_id 1009004
- Location of start within SEQ ID NO 19: at 48 nt.

40

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 13
- gi No. 3582320
- Description:
- % Identity: 32.6
- Alignment Length: 44
- Location of Alignment in SEQ ID NO 20: from 51 to 93

50

116

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 21
- Ceres seq\_id 1009005
- Location of start within SEQ ID NO 19: at 57 nt.

5

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 14
- gi No. 3582320
- Description:
- % Identity: 32.6
- Alignment Length: 44
- Location of Alignment in SEQ ID NO 21: from 48 to

15

90

## (Ba) Polypeptide Activities: Similar to Ring-H2 Zinc Finger Protein activities.

20

## Maximum Length Sequence:

related to:

Clone IDs:

7337

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 22
- Ceres seq\_id 1009345
- Alternative transcription start site(s) located in SEQ

30

ID NO 22:

2

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 23
- Ceres seq\_id 1009346
- Location of start within SEQ ID NO 22: at 50 nt.
- Location of Signal Peptide Cleavage Site within SEQ

ID NO 23: at 22 aa.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 15
- gi No. 3176705
- Description:
- % Identity: 53.3
- Alignment Length: 75
- Location of Alignment in SEQ ID NO 23: from 1 to

75

50

## (B) Polypeptide Sequence

289

117

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Pat. Appln. SEQ ID NO 24
- Ceres seq\_id 1009347
- Location of start within SEQ ID NO 22: at 62 nt.
- Location of Signal Peptide Cleavage Site within SEQ

5

ID NO 24: at 18 aa.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 16
- gi No. 3176705
- Description:
- % Identity: 53.3
- Alignment Length: 75
- Location of Alignment in SEQ ID NO 24: from 1 to

15

71

## (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot

20

specific gene.

## Maximum Length Sequence:

related to:

Clone IDs:

6349

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 25
- Ceres seq\_id 1010140
- Alternative transcription start site(s) located in SEQ

30

ID NO 25:

-31,-29,4,5,6,10,17,34,41,749

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 26
- Ceres seq\_id 1010141
- Location of start within SEQ ID NO 25: at 3 nt.

35

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Bacterial regulatory proteins, deoR family
- Location within SEQ ID NO 26: from 57 to 95 aa.

40

## (Dp) Related Amino Acid Sequences

- Alignment No. 17
- gi No. 3257798
- Description:
- % Identity: 40.6
- Alignment Length: 234
- Location of Alignment in SEQ ID NO 26: from 56 to

45

50

118

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 27  
 - Ceres seq\_id 1010142  
 - Location of start within SEQ ID NO 25: at 42 nt.

5

(C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)

- Bacterial regulatory proteins, deoR family  
 - Location within SEQ ID NO 27: from 44 to 82 aa.

10

(Dp) Related Amino Acid Sequences

- Alignment No. 18  
 - gi No. 3257798  
 - Description:  
 - % Identity: 40.6  
 - Alignment Length: 234  
 - Location of Alignment in SEQ ID NO 27: from 43 to

20

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 28  
 - Ceres seq\_id 1010143  
 - Location of start within SEQ ID NO 25: at 231 nt.

25

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences

- Alignment No. 19  
 - gi No. 3257798  
 - Description:  
 - % Identity: 40.6  
 - Alignment Length: 234  
 - Location of Alignment in SEQ ID NO 28: from 1 to

35

Maximum Length Sequence:

related to:

Clone IDs:

6261

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 29  
 - Ceres seq\_id 1010217  
 - Alternative transcription start site(s) located in SEQ  
 ID NO 29:  
 2,5,15

45

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 30  
 - Ceres seq\_id 1010218  
 - Location of start within SEQ ID NO 29: at 85 nt.

50

119

- Location of Signal Peptide Cleavage Site within SEQ  
 ID NO 30: at 22 aa.

- (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 20  
 - gi No. 3341723  
 - Description:  
 - % Identity: 64.3  
 - Alignment Length: 131  
 - Location of Alignment in SEQ ID NO 30: from 1 to

118

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 31  
 - Ceres seq\_id 1010219  
 - Location of start within SEQ ID NO 29: at 118 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ  
 ID NO 31: at 13 aa.

20

(C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 21  
 - gi No. 3341723  
 - Description:  
 - % Identity: 64.3  
 - Alignment Length: 131  
 - Location of Alignment in SEQ ID NO 31: from 1 to

30

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 32  
 - Ceres seq\_id 1010220  
 - Location of start within SEQ ID NO 29: at 121 nt.

35

(C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 22  
 - gi No. 3341723  
 - Description:  
 - % Identity: 64.3  
 - Alignment Length: 131  
 - Location of Alignment in SEQ ID NO 32: from 1 to

106

- (Ba) Polypeptide Activities: Similar to Constans like  
 Protein activities and Zinc Finger Protein Activities.

50



120

## Maximum Length Sequence:

related to:

Clone IDs:

6145

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 33

- Ceres seq\_id 1010302

- Alternative transcription start site(s) located in SEQ

ID NO 33:

-5,-3,-2,1,2,3,4,5,6,7,8,10,11,12,13,15,19,23,45,349

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 34

- Ceres seq\_id 1010303

- Location of start within SEQ ID NO 33: at 59 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Pathogenesis-related protein Bet v I family

- Location within SEQ ID NO 34: from 5 to 155 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 23

- gi No. 1321731

- Description:

- % Identity: 35.7

- Alignment Length: 159

- Location of Alignment in SEQ ID NO 34: from 5 to

30 155

## Maximum Length Sequence:

related to:

Clone IDs:

5180

Public Genomic DNA:

gi No: 4757410

Predicted Exons:

INTR 37202 ... 37397 OCKHAMG-CDNA

INTR 37493 ... 37825 OCKHAMG-CDNA

INIT 37271 ... 37397 OCKHAMG-CDS

TERM 37493 ... 37704 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 35

- Ceres seq\_id 1010815

- Alternative transcription start site(s) located in SEQ

ID NO 35:

15,16,17,18,19,29,31,34

(B) Polypeptide Sequence

121

- Pat. Appln. SEQ ID NO 36

- Ceres seq\_id 1010816

- Location of start within SEQ ID NO 35: at 70 nt.

5

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 24

- gi No. 2879811

- Description:

- % Identity: 88.4

- Alignment Length: 112

- Location of Alignment in SEQ ID NO 36: from 1 to

112

15

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 37

- Ceres seq\_id 1010817

- Location of start within SEQ ID NO 35: at 133 nt.

20

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 25

- gi No. 2879811

- Description:

- % Identity: 88.4

- Alignment Length: 112

- Location of Alignment in SEQ ID NO 37: from 1 to

30 91

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 38

- Ceres seq\_id 1010818

- Location of start within SEQ ID NO 35: at 257 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 38: at 26 aa.

(Ba) Polypeptide Activities: Similar to ribosomal protein

activities.

Maximum Length Sequence:

related to:

Clone IDs:

42842

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 39

- Ceres seq\_id 1011437

- Alternative transcription start site(s) located in SEQ

ID NO 39:

-28

122

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 40  
 - Ceres seq\_id 1011438  
 - Location of start within SEQ ID NO 39: at 2 nt.

5

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 41  
 - Ceres seq\_id 1011439  
 - Location of start within SEQ ID NO 39: at 1 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 41: at 29 aa.

10

## (C) Nomination and Annotation of Domains within

- Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences

- Alignment No. 26
- gi No. 3334271
- Description:
- % Identity: 29.6
- Alignment Length: 54
- Location of Alignment in SEQ ID NO 41: from 13 to

20

65

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 42  
 - Ceres seq\_id 1011440  
 - Location of start within SEQ ID NO 39: at 28 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 42: at 20 aa.

30

## (C) Nomination and Annotation of Domains within

- Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences

- Alignment No. 27
- gi No. 3334271
- Description:
- % Identity: 29.6
- Alignment Length: 54
- Location of Alignment in SEQ ID NO 42: from 4 to

40

56

- (Ba) Polypeptide Activities: Similar to NADH Oxidoreductase  
 MWFE Subunit Protein Activities.

45

Maximum Length Sequence:  
 related to:

- Clone IDs:  
 42475  
 (Ac) cDNA Polynucleotide Sequence

50

123

- Pat. Appln. SEQ ID NO 43
- Ceres seq\_id 1011616
- Alternative transcription start site(s) located in SEQ ID NO 43:

5

3,5,476

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 44  
 - Ceres seq\_id 1011617  
 - Location of start within SEQ ID NO 43: at 115 nt.

10

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 28
- gi No. 3859560
- Description:
- % Identity: 36.2
- Alignment Length: 225
- Location of Alignment in SEQ ID NO 44: from 31 to

20

248

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 45
- Ceres seq\_id 1011618
- Location of start within SEQ ID NO 43: at 406 nt.

25

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 29
- gi No. 3859560
- Description:
- % Identity: 36.2
- Alignment Length: 225
- Location of Alignment in SEQ ID NO 45: from 1 to

35

151

- (Ba) Polypeptide Activities: Similar to acyl-protein  
 thioesterases protein activities, calcium independent  
 phospholipase A2 activities, and carboxylesterase activities.

40

Maximum Length Sequence:  
 related to:

- Clone IDs:  
 42405  
 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 46  
 - Ceres seq\_id 1011631

50

124

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 47  
 - Ceres seq\_id 1011632  
 - Location of start within SEQ ID NO 46: at 3 nt.
- (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 30  
 - gi No. 3618318  
 - Description: 72.7  
 - % Identity: 72.7  
 - Alignment Length: 44  
 - Location of Alignment in SEQ ID NO 47: from 91 to 134
- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 48  
 - Ceres seq\_id 1011633  
 - Location of start within SEQ ID NO 46: at 9 nt.
- (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 31  
 - gi No. 3618318  
 - Description: 72.7  
 - % Identity: 72.7  
 - Alignment Length: 44  
 - Location of Alignment in SEQ ID NO 48: from 89 to 132
- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 49  
 - Ceres seq\_id 1011634  
 - Location of start within SEQ ID NO 46: at 15 nt.
- (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 32  
 - gi No. 3618318  
 - Description: 72.7  
 - % Identity: 72.7  
 - Alignment Length: 44  
 - Location of Alignment in SEQ ID NO 49: from 87 to 130

125

(Ba) Polypeptide Activities: Similar to Constans protein activities, and zinc finger protein activities.

- Maximum Length Sequence:  
 related to:  
 Clone IDs:  
 42240  
 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 50  
 - Ceres seq\_id 1011714
- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 51  
 - Ceres seq\_id 1011715  
 - Location of start within SEQ ID NO 50: at 2 nt.
- (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 - Bacterial mutT protein  
 - Location within SEQ ID NO 51: from 26 to 67 aa.
- (Dp) Related Amino Acid Sequences  
 - Alignment No. 33  
 - gi No. 2129134  
 - Description: 40.8  
 - % Identity: 40.8  
 - Alignment Length: 121  
 - Location of Alignment in SEQ ID NO 51: from 12 to 131
- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 52  
 - Ceres seq\_id 1011716  
 - Location of start within SEQ ID NO 50: at 14 nt.
- (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 - Bacterial mutT protein  
 - Location within SEQ ID NO 52: from 22 to 63 aa.
- (Dp) Related Amino Acid Sequences  
 - Alignment No. 34  
 - gi No. 2129134  
 - Description: 40.8  
 - % Identity: 40.8  
 - Alignment Length: 121  
 - Location of Alignment in SEQ ID NO 52: from 8 to 127
- (B) Polypeptide Sequence

- 126  
- Pat. Appln. SEQ ID NO 53  
- Ceres seq\_id 1011717  
- Location of start within SEQ ID NO 50: at 185 nt.
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 35  
- gi No. 2129134  
- Description:  
- % Identity: 40.8  
- Alignment Length: 121  
- Location of Alignment in SEQ ID NO 53: from 1 to
- 70
- Maximum Length Sequence:  
related to:  
Clone IDs:  
42169  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 54  
- Ceres seq\_id 1011784  
- Alternative transcription start site(s) located in SEQ  
ID NO 54:  
-15,-4,7,402
- 25
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 55  
- Ceres seq\_id 1011785  
- Location of start within SEQ ID NO 54: at 1 nt.  
- Location of Signal Peptide Cleavage Site within SEQ  
ID NO 55: at 42 aa.
- 30
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
- Plant lipid transfer protein family  
- Location within SEQ ID NO 55: from 45 to 108 aa.
- 35
- (Dp) Related Amino Acid Sequences  
- Alignment No. 36  
- gi No. 543565  
- Description:  
- % Identity: 56.5  
- Alignment Length: 85  
- Location of Alignment in SEQ ID NO 55: from 29 to
- 45
- 110
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 56  
- Ceres seq\_id 1011786  
- Location of start within SEQ ID NO 54: at 49 nt.
- 50

- 127  
- Location of Signal Peptide Cleavage Site within SEQ  
ID NO 56: at 26 aa.
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
- Plant lipid transfer protein family  
- Location within SEQ ID NO 56: from 29 to 92 aa.
- 5
- (Dp) Related Amino Acid Sequences  
- Alignment No. 37  
- gi No. 543565  
- Description:  
- % Identity: 56.5  
- Alignment Length: 85  
- Location of Alignment in SEQ ID NO 56: from 13 to
- 10
- 94
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 57  
- Ceres seq\_id 1011787  
- Location of start within SEQ ID NO 54: at 3 nt.
- 20
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences
- 25
- Maximum Length Sequence:  
related to:  
Clone IDs:  
41992  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 58  
- Ceres seq\_id 1011820  
- Alternative transcription start site(s) located in SEQ  
ID NO 58:  
-40,37
- 35
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 59  
- Ceres seq\_id 1011821  
- Location of start within SEQ ID NO 58: at 1 nt.
- 40
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 38  
- gi No. 3417418  
- Description:  
- % Identity: 23.6  
- Alignment Length: 207
- 45
- 50

- 128  
- Location of Alignment in SEQ ID NO 59: from 24 to 226
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 60  
- Ceres seq\_id 1011822  
- Location of start within SEQ ID NO 58: at 13 nt.
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 39  
- gi No. 3417418  
- Description: 23.6  
- % Identity: 23.6  
- Alignment Length: 207  
- Location of Alignment in SEQ ID NO 60: from 20 to 222
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 61  
- Ceres seq\_id 1011823  
- Location of start within SEQ ID NO 58: at 151 nt.
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 40  
- gi No. 3417418  
- Description: 23.6  
- % Identity: 23.6  
- Alignment Length: 207  
- Location of Alignment in SEQ ID NO 61: from 1 to 176
- (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot specific gene, plant specific gene.
- Maximum Length Sequence:  
related to:  
Clone IDs:  
41851  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 62  
- Ceres seq\_id 1011874
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 63

- 129  
- Ceres seq\_id 1011875  
- Location of start within SEQ ID NO 62: at 1 nt.
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 41  
- gi No. 2911044  
- Description: 78.7  
- % Identity: 78.7  
- Alignment Length: 95  
- Location of Alignment in SEQ ID NO 63: from 28 to 121
- (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot specific gene, plant specific gene.
- Maximum Length Sequence:  
related to:  
Clone IDs:  
41682  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 64  
- Ceres seq\_id 1011981
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 65  
- Ceres seq\_id 1011982  
- Location of start within SEQ ID NO 64: at 68 nt.
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 42  
- gi No. 4115355  
- Description: 100  
- % Identity: 100  
- Alignment Length: 52  
- Location of Alignment in SEQ ID NO 65: from 1 to 52
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 66  
- Ceres seq\_id 1011983  
- Location of start within SEQ ID NO 64: at 3 nt.
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
(B) Polypeptide Sequence

- 130  
- Pat. Appln. SEQ ID NO 67  
- Ceres seq\_id 1011984  
- Location of start within SEQ ID NO 64: at 483 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 67: at 19 aa.
- (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot specific gene, plant specific gene.
- 10
- Maximum Length Sequence:  
related to:  
Clone IDs:  
38470  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 68  
- Ceres seq\_id 1014547  
- Alternative transcription start site(s) located in SEQ ID NO 68:  
-39,-2,-1,2,3,7,8,9,10
- 15
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 69  
- Ceres seq\_id 1014548  
- Location of start within SEQ ID NO 68: at 1 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 69: at 32 aa.
- 20
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
- Plant lipid transfer protein family  
- Location within SEQ ID NO 69: from 44 to 110 aa.
- 25
- (Dp) Related Amino Acid Sequences  
- Alignment No. 43  
- gi No. 3062791  
- Description: 72.2  
- % Identity: 72.2  
- Alignment Length: 90  
- Location of Alignment in SEQ ID NO 69: from 21 to 110
- 30
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 70  
- Ceres seq\_id 1014549  
- Location of start within SEQ ID NO 68: at 25 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 70: at 24 aa.
- 35
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 73  
- Ceres seq\_id 1014997  
- Location of start within SEQ ID NO 71: at 65 nt.
- 40
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- 45
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- 50

- 131  
(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
- Plant lipid transfer protein family  
- Location within SEQ ID NO 70: from 36 to 102 aa.
- 5
- (Dp) Related Amino Acid Sequences  
- Alignment No. 44  
- gi No. 3062791  
- Description: 72.2  
- % Identity: 72.2  
- Alignment Length: 90  
- Location of Alignment in SEQ ID NO 70: from 13 to 102
- 10
- Maximum Length Sequence:  
related to:  
Clone IDs:  
38004  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 71  
- Ceres seq\_id 1014995  
- Alternative transcription start site(s) located in SEQ ID NO 71:  
2,3,4
- 15
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 72  
- Ceres seq\_id 1014996  
- Location of start within SEQ ID NO 71: at 2 nt.
- 20
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
- Aminotransferase class IV  
- Location within SEQ ID NO 72: from 71 to 334 aa.
- 25
- (Dp) Related Amino Acid Sequences  
- Alignment No. 45  
- gi No. 3540183  
- Description: 54.7  
- % Identity: 54.7  
- Alignment Length: 287  
- Location of Alignment in SEQ ID NO 72: from 56 to 341
- 30
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 73  
- Ceres seq\_id 1014997  
- Location of start within SEQ ID NO 71: at 65 nt.
- 35
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- 40
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- 45
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- 50

132

- Aminotransferase class IV
- Location within SEQ ID NO 73: from 50 to 313 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 46
- gi No. 3540183
- Description:
- % Identity: 54.7
- Alignment Length: 287
- Location of Alignment in SEQ ID NO 73: from 35 to 320

10

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 74
- Ceres seq\_id 1014998
- Location of start within SEQ ID NO 71: at 167 nt.

15

## (C) Nomination and Annotation of Domains within

- Predicted Polypeptide(s)
- Aminotransferase class IV
- Location within SEQ ID NO 74: from 16 to 279 aa.

20

## (Dp) Related Amino Acid Sequences

- Alignment No. 47
- gi No. 3540183
- Description:
- % Identity: 54.7
- Alignment Length: 287
- Location of Alignment in SEQ ID NO 74: from 1 to 286

30

## Maximum Length Sequence:

related to:

Clone IDs:

37701

Public Genomic DNA:

gi No: 4699904

Predicted Exons:

SINGLE 38530 ... 37988 OCKHAMG-CDS

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 75
- Ceres seq\_id 1015323
- Alternative transcription start site(s) located in SEQ ID NO 75:

-1,5,6,12

45

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 76
- Ceres seq\_id 1015324
- Location of start within SEQ ID NO 75: at 59 nt.

50

133

- Location of Signal Peptide Cleavage Site within SEQ ID NO 76: at 25 aa.

## (C) Nomination and Annotation of Domains within

## (Dp) Related Amino Acid Sequences

- Alignment No. 48
- gi No. 3860308
- Description:
- % Identity: 44.5
- Alignment Length: 140
- Location of Alignment in SEQ ID NO 76: from 56 to 175

10

- (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot

specific gene, plant

specific gene.

## Maximum Length Sequence:

related to:

Clone IDs:

364

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 77
- Ceres seq\_id 1016486
- Alternative transcription start site(s) located in SEQ ID NO 77:

17,19,20,21,22,23,29,35,38

30

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 78
- Ceres seq\_id 1016487
- Location of start within SEQ ID NO 77: at 1 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 78: at 46 aa.

35

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Gamma-thionins family
- Location within SEQ ID NO 78: from 55 to 101 aa.

40

## (Dp) Related Amino Acid Sequences

- Alignment No. 49
- gi No. 4038039
- Description:
- % Identity: 100
- Alignment Length: 77
- Location of Alignment in SEQ ID NO 78: from 25 to 101

50

134

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 79
- Ceres seq\_id 1016488
- Location of start within SEQ ID NO 77: at 73 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 79: at 22 aa.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Gamma-thionins family
- Location within SEQ ID NO 79: from 31 to 77 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 50
- gi No. 4038039
- Description:
- % Identity: 100
- Alignment Length: 77
- Location of Alignment in SEQ ID NO 79: from 1 to

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 80
- Ceres seq\_id 1016489
- Location of start within SEQ ID NO 77: at 118 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Gamma-thionins family
- Location within SEQ ID NO 80: from 16 to 62 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 51
- gi No. 4038039
- Description:
- % Identity: 100
- Alignment Length: 77
- Location of Alignment in SEQ ID NO 80: from 1 to

## Maximum Length Sequence:

related to:  
Clone IDs:

33891

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 81
- Ceres seq\_id 1018341
- Alternative transcription start site(s) located in SEQ ID NO 81:

4

135

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 82
- Ceres seq\_id 1018342
- Location of start within SEQ ID NO 81: at 71 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 52
- gi No. 1064887
- Description:
- % Identity: 60
- Alignment Length: 65
- Location of Alignment in SEQ ID NO 82: from 2 to

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 83
- Ceres seq\_id 1018343
- Location of start within SEQ ID NO 81: at 143 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 53
- gi No. 1064887
- Description:
- % Identity: 60
- Alignment Length: 65
- Location of Alignment in SEQ ID NO 83: from 1 to

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 84
- Ceres seq\_id 1018344
- Location of start within SEQ ID NO 81: at 146 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 54
- gi No. 1064887
- Description:
- % Identity: 60
- Alignment Length: 65
- Location of Alignment in SEQ ID NO 84: from 1 to

- (Ba) Polypeptide Activities: Similar to pollen coat protein activities and IEA protein activities.



136

## Maximum Length Sequence:

related to:  
Clone IDs:

33828

- 5 (Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 85  
- Ceres seq\_id 1018382  
- Alternative transcription start site(s) located in SEQ ID NO 85:

10 2,4,5,6,7,8,9,10,11,14

## (B) Polypeptide Sequence

- 15 - Pat. Appln. SEQ ID NO 86  
- Ceres seq\_id 1018383  
- Location of start within SEQ ID NO 85: at 22 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 86: at 24 aa.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- 20 - Plant lipid transfer protein family  
- Location within SEQ ID NO 86: from 28 to 115 aa.

## (Dp) Related Amino Acid Sequences

- 25 - Alignment No. 55  
- gi No. 899224  
- Description:  
- % Identity: 78.2  
- Alignment Length: 119  
- Location of Alignment in SEQ ID NO 86: from 1 to

119

## (B) Polypeptide Sequence

- 35 - Pat. Appln. SEQ ID NO 87  
- Ceres seq\_id 1018384  
- Location of start within SEQ ID NO 85: at 73 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- 40 - Plant lipid transfer protein family  
- Location within SEQ ID NO 87: from 11 to 98 aa.

## (Dp) Related Amino Acid Sequences

- 45 - Alignment No. 56  
- gi No. 899224  
- Description:  
- % Identity: 78.2  
- Alignment Length: 119  
- Location of Alignment in SEQ ID NO 87: from 1 to

50 102

137

## (B) Polypeptide Sequence

- 5 - Pat. Appln. SEQ ID NO 88  
- Ceres seq\_id 1018385  
- Location of start within SEQ ID NO 85: at 3 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

10 Maximum Length Sequence:  
related to:

Clone IDs:

30349

## (Ac) cDNA Polynucleotide Sequence

- 15 - Pat. Appln. SEQ ID NO 89  
- Ceres seq\_id 1020666  
- Alternative transcription start site(s) located in SEQ ID NO 89:

33,35,39,40,42,43,44,45,64,173

## (B) Polypeptide Sequence

- 20 - Pat. Appln. SEQ ID NO 90  
- Ceres seq\_id 1020667  
- Location of start within SEQ ID NO 89: at 118 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- 25 - Ribosomal protein S7e  
- Location within SEQ ID NO 90: from 7 to 187 aa.

## (Dp) Related Amino Acid Sequences

- 30 - Alignment No. 57  
- gi No. 3851636  
- Description:  
- % Identity: 77.4  
- Alignment Length: 190  
- Location of Alignment in SEQ ID NO 90: from 1 to

190

## (B) Polypeptide Sequence

- 40 - Pat. Appln. SEQ ID NO 91  
- Ceres seq\_id 1020668  
- Location of start within SEQ ID NO 89: at 271 nt.

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- 45 - Ribosomal protein S7e  
- Location within SEQ ID NO 91: from 1 to 136 aa.

## (Dp) Related Amino Acid Sequences

- 50 - Alignment No. 58

- 138
- gi No. 3851636
  - Description:
  - % Identity: 77.4
  - Alignment Length: 190
  - Location of Alignment in SEQ ID NO 91: from 1 to
- 5 139
- Maximum Length Sequence:  
related to:  
Clone IDs:  
30113
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 92
  - Ceres seq\_id 1020784
- 15
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 93
  - Ceres seq\_id 1020785
  - Location of start within SEQ ID NO 92: at 60 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 93: at 25 aa.
- 20
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences
- Alignment No. 59
  - gi No. 3860308
  - Description:
  - % Identity: 44.5
  - Alignment Length: 140
  - Location of Alignment in SEQ ID NO 93: from 56 to
- 30 175
- (Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot specific gene, plant specific gene.
- 35
- Maximum Length Sequence:  
related to:  
Clone IDs:  
29120
- Public Genomic DNA:  
gi No: 5822667
- Predicted Exons:
- INTR 68772 ... 69532 OCKHAMG-CDNA
- 45
- SINGLE 68846 ... 69325 OCKHAMG-CDS
- gi No: 6041831
- Predicted Exons:
- 50

- 139
- INTR 63702 ... 64462 OCKHAMG-CDNA
- SINGLE 63776 ... 64255 OCKHAMG-CDS
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 94
  - Ceres seq\_id 1021525
  - Alternative transcription start site(s) located in SEQ ID NO 94:
- 25, 26, 27, 28, 29, 35, 36, 39, 51, 53, 54, 68
- 10
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 95
  - Ceres seq\_id 1021526
  - Location of start within SEQ ID NO 94: at 75 nt.
- 15
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences
- Alignment No. 60
  - gi No. 4388980
  - Description:
  - % Identity: 29.1
  - Alignment Length: 110
  - Location of Alignment in SEQ ID NO 95: from 47 to
- 20 156
- (Ba) Polypeptide Activities: Similar to adrenodoxin precursor  
protein activities and adrenal ferredoxin activities.
- 30
- Maximum Length Sequence:  
related to:  
Clone IDs:  
2891
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 96
  - Ceres seq\_id 1021563
  - Alternative transcription start site(s) located in SEQ ID NO 96:
- 16, 28, 29, 30, 31, 35, 36, 43, 74, 77, 80, 88, 89, 90, 95
- 40
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 97
  - Ceres seq\_id 1021564
  - Location of start within SEQ ID NO 96: at 2 nt.
- 45
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences
- Alignment No. 61
- 50

140

- gi No. 2493089
- Description:
- % Identity: 29.7
- Alignment Length: 77
- Location of Alignment in SEQ ID NO 97: from 52 to

5

125

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 98
- Ceres seq\_id 1021565
- Location of start within SEQ ID NO 96: at 116 nt.

10

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 62
- gi No. 2493089
- Description:
- % Identity: 29.7
- Alignment Length: 77
- Location of Alignment in SEQ ID NO 98: from 14 to

20

87

- (Ba) Polypeptide Activities: Similar to ATPK-mouse activities, ATP synthase activities, and mitochondrial F-Chain activities.

25

## Maximum Length Sequence:

related to:

Clone IDs:

28979

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 99
- Ceres seq\_id 1021576
- Alternative transcription start site(s) located in SEQ

35

ID NO 99:

-13,-

11,3,5,7,11,12,13,15,24,26,28,29,31,34,35,36,37,38,39

- 43,44,45,46,47,48,50,51,52,56,58,59,61,67,68,70,75,82,83,91,1

40

12

244,313,318

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 100
- Ceres seq\_id 1021577
- Location of start within SEQ ID NO 99: at 109 nt.

45

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## - Photosystem I psaG / psaK

50

141

- Location within SEQ ID NO 100: from 50 to 130 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 63
- gi No. 3885511
- Description:
- % Identity: 81.1
- Alignment Length: 128
- Location of Alignment in SEQ ID NO 100: from 1 to

5

10

127

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 101
- Ceres seq\_id 1021578
- Location of start within SEQ ID NO 99: at 121 nt.

15

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## - Photosystem I psaG / psaK

- Location within SEQ ID NO 101: from 46 to 126 aa.

20

## (Dp) Related Amino Acid Sequences

- Alignment No. 64
- gi No. 3885511
- Description:
- % Identity: 81.1
- Alignment Length: 128
- Location of Alignment in SEQ ID NO 101: from 1 to

123

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 102
- Ceres seq\_id 1021579
- Location of start within SEQ ID NO 99: at 124 nt.

30

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## - Photosystem I psaG / psaK

- Location within SEQ ID NO 102: from 45 to 125 aa.

35

40

## (Dp) Related Amino Acid Sequences

- Alignment No. 65
- gi No. 3885511
- Description:
- % Identity: 81.1
- Alignment Length: 128
- Location of Alignment in SEQ ID NO 102: from 1 to

45

122

## Maximum Length Sequence:

related to:

50

142

## Clone IDs:

28177

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 103

- Ceres seq\_id 1021927

5 - Alternative transcription start site(s) located in SEQ

ID NO 103:

2,12,20,23,29,36,46,47,53

10

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 104

- Ceres seq\_id 1021928

- Location of start within SEQ ID NO 103: at 67 nt.

15

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 66

- gi No. 4263779

- Description:

- % Identity: 28.2

- Alignment Length: 177

- Location of Alignment in SEQ ID NO 104: from 14

to 182

25

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 105

- Ceres seq\_id 1021929

- Location of start within SEQ ID NO 103: at 172 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 105: at 23 aa.

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 67

- gi No. 4263779

- Description:

- % Identity: 28.2

- Alignment Length: 177

- Location of Alignment in SEQ ID NO 105: from 1 to

147

45

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 106

- Ceres seq\_id 1021930

- Location of start within SEQ ID NO 103: at 220 nt.

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

143

- Alignment No. 68

- gi No. 4263779

- Description:

- % Identity: 28.2

- Alignment Length: 177

- Location of Alignment in SEQ ID NO 106: from 1 to

131

## (Ba) Polypeptide Activities: Arabidopsis specific gene,

dicot specific gene, plant specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

2807

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 107

- Ceres seq\_id 1021945

20

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 108

- Ceres seq\_id 1021946

- Location of start within SEQ ID NO 107: at 1 nt.

25

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 69

- gi No. 3334132

- Description:

- % Identity: 27.4

- Alignment Length: 114

- Location of Alignment in SEQ ID NO 108: from 28

35 to 136

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 109

- Ceres seq\_id 1021947

- Location of start within SEQ ID NO 107: at 163 nt.

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 70

- gi No. 3334132

- Description:

- % Identity: 27.4

- Alignment Length: 114

- Location of Alignment in SEQ ID NO 109: from 1 to

82

144

(Ba) Polypeptide Activities: Similar to C21 ORF4 - membrane protein activities

5 Maximum Length Sequence:

related to:

Clone IDs:

27792

(Ac) cDNA Polynucleotide Sequence

10 - Pat. Appln. SEQ ID NO 110

- Ceres seq\_id 1022170

- Alternative transcription start site(s) located in SEQ

ID NO 110:

-4,-1,32,68

15 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 111

- Ceres seq\_id 1022171

- Location of start within SEQ ID NO 110: at 92 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 71

- gi No. 1173456

- Description:

- % Identity: 54.7

- Alignment Length: 129

- Location of Alignment in SEQ ID NO 111: from 4 to

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 112

- Ceres seq\_id 1022172

- Location of start within SEQ ID NO 110: at 191 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 72

- gi No. 1173456

- Description:

- % Identity: 54.7

- Alignment Length: 129

- Location of Alignment in SEQ ID NO 112: from 1 to

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 113

- Ceres seq\_id 1022173

- Location of start within SEQ ID NO 110: at 1 nt.

SUBSTITUTE SHEET (RULE 26)

145

- Location of Signal Peptide Cleavage Site within SEQ ID NO 113: at 19 aa.

5 (Ba) Polypeptide Activities: Similar to small nuclear ribonucleoprotein activities.

Maximum Length Sequence:

related to:

Clone IDs:

27167

Public Genomic DNA:

gi No: 3046850

Predicted Exons:

INIT 45217 ... 45131 OCKHAMG-CDS

INTR 44695 ... 44629 OCKHAMG-CDS

TERM 44554 ... 44286 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 114

- Ceres seq\_id 1022554

- Alternative transcription start site(s) located in SEQ

ID NO 114:

-49,-

19,7,22,23,28,29,30,31,32,33,35,36,37,51,57,67,75,81,82

85

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 115

- Ceres seq\_id 1022555

- Location of start within SEQ ID NO 114: at 107 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Heme-binding domain in cytochrome b5 and

oxidoreductases

- Location within SEQ ID NO 115: from 7 to 84 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 73

- gi No. 4240122

- Description:

- % Identity: 100

- Alignment Length: 140

- Location of Alignment in SEQ ID NO 115: from 1 to

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 116

- Ceres seq\_id 1022556

- Location of start within SEQ ID NO 114: at 317 nt.

SUBSTITUTE SHEET (RULE 26)

146

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 74
- Description: 100
- % Identity: 100
- Alignment Length: 140
- Location of Alignment in SEQ ID NO 116: from 1 to 70

## Maximum Length Sequence:

## related to:

## Clone IDs:

27109

## Public Genomic DNA:

gi No: 6449507

## Predicted Exons:

INIT 94711 ... 94519 OCKHANG-CDS  
 INTR 94417 ... 94326 OCKHANG-CDS  
 INTR 94249 ... 94131 OCKHANG-CDS  
 TERM 94046 ... 93968 OCKHANG-CDS

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 117
- Ceres seq\_id 1022594
- Alternative transcription start site(s) located in SEQ ID NO 117: 2,15,24,25,66,69,72,74

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 118
- Ceres seq\_id 1022595
- Location of start within SEQ ID NO 117: at 1 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Universal stress protein family
- Location within SEQ ID NO 118: from 98 to 183 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 75
- gi No. 2160182
- Description: 37.5
- % Identity: 37.5
- Alignment Length: 160
- Location of Alignment in SEQ ID NO 118: from 39 to 186

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 119
- Ceres seq\_id 1022596

147

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Universal stress protein family
- Location within SEQ ID NO 119: from 71 to 156 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 76
- gi No. 2160182
- Description: 37.5
- % Identity: 37.5
- Alignment Length: 160
- Location of Alignment in SEQ ID NO 119: from 12 to 159

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 120
- Ceres seq\_id 1022597
- Location of start within SEQ ID NO 117: at 106 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Universal stress protein family
- Location within SEQ ID NO 120: from 63 to 148 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 77
- gi No. 2160182
- Description: 37.5
- % Identity: 37.5
- Alignment Length: 160
- Location of Alignment in SEQ ID NO 120: from 4 to 151

## (Ba) Polypeptide Activities: Similar to protein in

methanobacterium thermoautotrophicum activities.

## Maximum Length Sequence:

## related to:

## Clone IDs:

26994

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 121
- Ceres seq\_id 1022621
- Alternative transcription start site(s) located in SEQ ID NO 121: 2,7,9,13,35,38,45,57

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 122

148

- Ceres seq\_id 1022622
- Location of start within SEQ ID NO 121: at 86 nt.

## (C) Nomination and Annotation of Domains within

## 5 Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 78
- gi No. 3256599
- Description:
- % Identity: 32
- Alignment length: 128
- Location of Alignment in SEQ ID NO 122: from 7 to 128

- 15 (Ba) Polypeptide Activities: Similar to structural cell wall protein activities, and larval gene protein in the Fruit fly activities.

## 20 Maximum Length Sequence:

related to:

Clone IDs:

23518

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 123
- Ceres seq\_id 1024375
- Alternative transcription start site(s) located in SEQ ID NO 123: 7,8,10,17,18,24,35,41,42,43,44,46,48,52,54,59

## 30 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 124
- Ceres seq\_id 1024376
- Location of start within SEQ ID NO 123: at 130 nt.

## (C) Nomination and Annotation of Domains within

## 35 Predicted Polypeptide(s)

## - Ribosomal protein S7e

- Location within SEQ ID NO 124: from 7 to 187 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 79
- gi No. 3851636
- Description:
- % Identity: 76.6
- Alignment length: 188
- Location of Alignment in SEQ ID NO 124: from 1 to 188

## 50 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 125

149

- Ceres seq\_id 1024377
- Location of start within SEQ ID NO 123: at 283 nt.

## (C) Nomination and Annotation of Domains within

## 5 Predicted Polypeptide(s)

## - Ribosomal protein S7e

- Location within SEQ ID NO 125: from 1 to 136 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 80
- gi No. 3851636
- Description:
- % Identity: 76.6
- Alignment length: 188
- Location of Alignment in SEQ ID NO 125: from 1 to 137

## Maximum Length Sequence:

related to:

Clone IDs:

23170

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 126
- Ceres seq\_id 1024535
- Alternative transcription start site(s) located in SEQ ID NO 126: 2,8,11,31,46,47,48

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 127
- Ceres seq\_id 1024536
- Location of start within SEQ ID NO 126: at 115 nt.

## (C) Nomination and Annotation of Domains within

## 35 Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 81
- gi No. 2621731
- Description:
- % Identity: 35.2
- Alignment length: 88
- Location of Alignment in SEQ ID NO 127: from 5 to 92

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 128
- Ceres seq\_id 1024537
- Location of start within SEQ ID NO 126: at 3 nt.

## 50 (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences  
150

- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 129  
- Ceres seq\_id 1024538  
5 - Location of start within SEQ ID NO 126: at 253 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- (Dp) Related Amino Acid Sequences

- Alignment No. 82  
- gi No. 2621731  
- Description:  
- % Identity: 35.2  
- Alignment Length: 88  
15 - Location of Alignment in SEQ ID NO 129: from 1 to  
46

- (Ba) Polypeptide Activities: Similar to small nuclear  
20 ribonucleoprotein activities.

Maximum Length Sequence:

related to:

Clone IDs:

21228

Public Genomic DNA:

gi No: 4539402

Predicted Exons:

- INIT 36726 ... 36352 GENBANK  
INTR 36257 ... 36135 GENBANK  
TERM 35555 ... 35469 GENBANK

- INTR 36791 ... 36352 OCKHAMG-CDNA  
INTR 36257 ... 36135 OCKHAMG-CDNA  
INTR 35555 ... 35325 OCKHAMG-CDNA

- INIT 36726 ... 36352 OCKHAMG-CDS  
INTR 36257 ... 36135 OCKHAMG-CDS  
TERM 35555 ... 35469 OCKHAMG-CDS  
40 (Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 130  
- Ceres seq\_id 1025683  
- Alternative transcription start site(s) located in SEQ

45 ID NO 130:

2,8,22

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 131  
- Ceres seq\_id 1025684  
50 - Location of start within SEQ ID NO 130: at 3 nt.

151

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Ribosomal protein L6  
5 - Location within SEQ ID NO 131: from 34 to 216 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 83  
- gi No. 266945  
- Description:  
- % Identity: 84  
10 - Alignment Length: 194  
- Location of Alignment in SEQ ID NO 131: from 23  
to 216

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 132  
- Ceres seq\_id 1025685  
20 - Location of start within SEQ ID NO 130: at 69 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Ribosomal protein L6  
- Location within SEQ ID NO 132: from 12 to 194 aa.  
25

(Dp) Related Amino Acid Sequences

- Alignment No. 84  
- gi No. 266945  
- Description:  
- % Identity: 84  
30 - Alignment Length: 194  
- Location of Alignment in SEQ ID NO 132: from 1 to  
194

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 133  
- Ceres seq\_id 1025686  
35 - Location of start within SEQ ID NO 130: at 96 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Ribosomal protein L6  
40 - Location within SEQ ID NO 133: from 3 to 185 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 85  
- gi No. 266945  
- Description:  
- % Identity: 84  
45 - Alignment Length: 194



- 185
- 152
- Location of Alignment in SEQ ID NO 133: from 1 to
- 5
- Maximum Length Sequence:  
related to:  
Clone IDs:  
19274
- (Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 134  
- Ceres seq\_id 1027152  
- Alternative transcription start site(s) located in SEQ ID NO 134:  
-350,3,4,10,11,13,222
- 10
- 15
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 135  
- Ceres seq\_id 1027153  
- Location of start within SEQ ID NO 134: at 3 nt.
- 20
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 86  
- gi No. 2879811  
- Description:  
- % Identity: 86.6  
- Alignment Length: 112  
- Location of Alignment in SEQ ID NO 135: from 26 to 137
- 25
- 30
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 136  
- Ceres seq\_id 1027154  
- Location of start within SEQ ID NO 134: at 78 nt.
- 35
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 87  
- gi No. 2879811  
- Description:  
- % Identity: 86.6  
- Alignment Length: 112  
- Location of Alignment in SEQ ID NO 136: from 1 to
- 40
- 45
- 112
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 137  
- Ceres seq\_id 1027155  
- Location of start within SEQ ID NO 134: at 141 nt.
- 50

- 153
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 88  
- gi No. 2879811  
- Description:  
- % Identity: 86.6  
- Alignment Length: 112  
- Location of Alignment in SEQ ID NO 137: from 1 to
- 5
- 10
- 91
- (Ba) Polypeptide Activities: Similar to ribosomal L30 protein activities.
- 15
- Maximum Length Sequence:  
related to:  
Clone IDs:  
17835
- 20
- (Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 138  
- Ceres seq\_id 1028095  
- Alternative transcription start site(s) located in SEQ ID NO 138:  
-2,2,3,4,5,6,12,14,18,22,26,40,42,44,45,46,47
- 25
- 30
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 139  
- Ceres seq\_id 1028096  
- Location of start within SEQ ID NO 138: at 55 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 139: at 29 aa.
- 35
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 89  
- gi No. 4336325  
- Description:  
- % Identity: 31.7  
- Alignment Length: 126  
- Location of Alignment in SEQ ID NO 139: from 15 to 135
- 40
- 45
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 140  
- Ceres seq\_id 1028097  
- Location of start within SEQ ID NO 138: at 214 nt.
- 50
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- 154  
(Dp) Related Amino Acid Sequences  
- Alignment No. 90  
- gi No. 4336325  
- Description:  
- % Identity: 31.7  
- Alignment Length: 126  
- Location of Alignment in SEQ ID NO 140: from 1 to 82
- 5
- 10 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 141  
- Ceres seq\_id 1028098  
- Location of start within SEQ ID NO 138: at 3 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 141: at 16 aa.
- 15 (Ba) Polypeptide Activities: Similar to human C214 membrane protein activities.
- 20 Maximum Length Sequence:  
related to:  
Clone IDs:  
17075  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 142  
- Ceres seq\_id 1028608  
- Alternative transcription start site(s) located in SEQ ID NO 142:  
-4,2,28,31,36,49,59
- 25 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 143  
- Ceres seq\_id 1028609  
- Location of start within SEQ ID NO 142: at 95 nt.
- 35 (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 91  
- gi No. 2735528  
- Description:  
- % Identity: 33.9  
- Alignment Length: 118  
- Location of Alignment in SEQ ID NO 143: from 64 to 178
- 45 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 144  
- Ceres seq\_id 1028610  
- Location of start within SEQ ID NO 142: at 176 nt.
- 50

155

- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 92  
- gi No. 2735528  
- Description:  
- % Identity: 33.9  
- Alignment Length: 118  
- Location of Alignment in SEQ ID NO 144: from 37 to 151
- 10
- 15 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 145  
- Ceres seq\_id 1028611  
- Location of start within SEQ ID NO 142: at 381 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 145: at 41 aa.
- 20 (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
Maximum Length Sequence:  
related to:  
Clone IDs:  
1505  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 146  
- Ceres seq\_id 1030069
- 25
- 30 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 147  
- Ceres seq\_id 1030070  
- Location of start within SEQ ID NO 146: at 2 nt.
- 35 (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 93  
- gi No. 4335755  
- Description:  
- % Identity: 63.6  
- Alignment Length: 143  
- Location of Alignment in SEQ ID NO 147: from 34 to 171
- 40
- 45 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 148  
- Ceres seq\_id 1030071
- 50

156

- Location of start within SEQ ID NO 146: at 50 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 94
- gi No. 4335755
- Description:
- % Identity: 63.6
- Alignment Length: 143
- Location of Alignment in SEQ ID NO 148: from 18 to 155

10

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 149
- Ceres seq\_id 1030072
- Location of start within SEQ ID NO 146: at 170 nt.

15

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 95
- gi No. 4335755
- Description:
- % Identity: 63.6
- Alignment Length: 143
- Location of Alignment in SEQ ID NO 149: from 1 to 115

25

(Ba) Polypeptide Activities: Similar to hydroxyproline-rich protein activities.

30

Maximum Length Sequence:

related to:

Clone IDs:

12487

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 150
- Ceres seq\_id 1032069
- Alternative transcription start site(s) located in SEQ ID NO 150: 4,17

35

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 151
- Ceres seq\_id 1032070
- Location of start within SEQ ID NO 150: at 74 nt.

45

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

50

157

(Dp) Related Amino Acid Sequences

- Alignment No. 96
- gi No. 3386621
- Description:
- % Identity: 97.4
- Alignment Length: 234
- Location of Alignment in SEQ ID NO 151: from 1 to 231

5

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 152
- Ceres seq\_id 1032071
- Location of start within SEQ ID NO 150: at 122 nt.

10

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 97
- gi No. 3386621
- Description:
- % Identity: 97.4
- Alignment Length: 234
- Location of Alignment in SEQ ID NO 152: from 1 to 215

20

25

(Ba) Polypeptide Activities: Arabidopsis specific gene, dicot specific gene, plant specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

11466

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 153
- Ceres seq\_id 1033557
- Alternative transcription start site(s) located in SEQ ID NO 153: 62,64,65,67,72,73,74,75,166

35

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 154
- Ceres seq\_id 1033558
- Location of start within SEQ ID NO 153: at 94 nt.

40

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Ribosomal protein L34
- Location within SEQ ID NO 154: from 105 to 145

45

aa.

(Dp) Related Amino Acid Sequences

50

- 158
- Alignment No. 98
  - gi No. 132909
  - Description:
  - % Identity: 70
  - Alignment Length: 30
  - Location of Alignment in SEQ ID NO 154: from 116 to 145
- 5
- (Ba) Polypeptide Activities: Similar to 50S ribosomal protein I34 activities.
- 10
- Maximum Length Sequence:  
related to:  
Clone IDs:  
21589  
106951
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 155
  - Ceres seq\_id 1034688
  - Alternative transcription start site(s) located in SEQ ID NO 155:
- 20
- Clone 21589 starts at 2 and ends at in cDNA.
- 25
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 156
  - Ceres seq\_id 1034689
  - Location of start within SEQ ID NO 155: at 2 nt.
- 30
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences
- Alignment No. 99
  - gi No. 4335755
  - Description:
  - % Identity: 65
  - Alignment Length: 143
  - Location of Alignment in SEQ ID NO 156: from 33 to 170
- 35
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 157
  - Ceres seq\_id 1034690
  - Location of start within SEQ ID NO 155: at 47 nt.
- 40
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences
- Alignment No. 100
  - gi No. 4335755
  - Description:
- 50

- 159
- % Identity: 65
  - Alignment Length: 143
  - Location of Alignment in SEQ ID NO 157: from 18 to 155
- 5
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 158
  - Ceres seq\_id 1034691
  - Location of start within SEQ ID NO 155: at 167 nt.
- 10
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences
- Alignment No. 101
  - gi No. 4335755
  - Description:
  - % Identity: 65
  - Alignment Length: 143
  - Location of Alignment in SEQ ID NO 158: from 1 to 115
- 20
- (Ba) Polypeptide Activities: Similar to hydroxy proline rich glycoprotein activities.
- 25
- Maximum Length Sequence:  
related to:  
Clone IDs:  
10433
- 30
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 159
  - Ceres seq\_id 1035033
  - Alternative transcription start site(s) located in SEQ ID NO 159:
- 35
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 160
  - Ceres seq\_id 1035034
  - Location of start within SEQ ID NO 159: at 3 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 160: at 47 aa.
- 40
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences
- Alignment No. 102
  - gi No. 3062795
  - Description:
  - % Identity: 39.8
  - Alignment Length: 83
- 50

- 160  
- Location of Alignment in SEQ ID NO 160: from 24 to 99
- 5 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 161  
- Ceres seq\_id 1035035  
- Location of start within SEQ ID NO 159: at 69 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 161: at 25 aa.
- 10 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 103  
- gi No. 3062795  
- Description:  
- % Identity: 39.8  
- Alignment Length: 83  
- Location of Alignment in SEQ ID NO 161: from 2 to 77
- 15 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 162  
- Ceres seq\_id 1035036  
- Location of start within SEQ ID NO 159: at 72 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 162: at 24 aa.
- 25 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 104  
- gi No. 3062795  
- Description:  
- % Identity: 39.8  
- Alignment Length: 83  
- Location of Alignment in SEQ ID NO 162: from 1 to 76
- 30 (Ba) Polypeptide Activities: Similar to Pollen coat protein activities.
- 35 Maximum Length Sequence:  
related to:  
Clone IDs:  
10394  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 163  
- Ceres seq\_id 1035071
- 40 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 166  
- Ceres seq\_id 1035074  
- Location of start within SEQ ID NO 163: at 513 nt.
- 45 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 107
- 50 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 164  
- Ceres seq\_id 1035072  
- Location of start within SEQ ID NO 163: at 66 nt.
- 161  
- Alternative transcription start site(s) located in SEQ ID NO 163:  
-2,-1,2,3,18,19,21,24,25,56,66
- 5 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 164  
- Ceres seq\_id 1035072  
- Location of start within SEQ ID NO 163: at 66 nt.
- 10 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
- KH domain  
- Location within SEQ ID NO 164: from 47 to 95 aa.
- 15 (Dp) Related Amino Acid Sequences  
- Alignment No. 105  
- gi No. 133940  
- Description:  
- % Identity: 76  
- Alignment Length: 246  
- Location of Alignment in SEQ ID NO 164: from 1 to 148
- 20 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 165  
- Ceres seq\_id 1035073  
- Location of start within SEQ ID NO 163: at 195 nt.
- 25 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
- KH domain  
- Location within SEQ ID NO 165: from 4 to 52 aa.
- 30 (Dp) Related Amino Acid Sequences  
- Alignment No. 106  
- gi No. 133940  
- Description:  
- % Identity: 76  
- Alignment Length: 246  
- Location of Alignment in SEQ ID NO 165: from 1 to 105
- 35 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 166  
- Ceres seq\_id 1035074  
- Location of start within SEQ ID NO 163: at 513 nt.
- 40 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 107

162

- gi No. 133940
- Description:
- % Identity: 76
- Alignment Length: 246
- Location of Alignment in SEQ ID NO 166: from 1 to

5 97

## Maximum Length Sequence:

related to:

Clone IDs:

10511

## Public Genomic DNA:

gi No: 4539290

Predicted Exons:

INTR 5588 ... 5314 OCKHAMG-CDNA

15

INTR 5517 ... 5314 OCKHAMG-CDNA

gi No: 4914454

Predicted Exons:

INTR 38937 ... 38663 OCKHAMG-CDNA

20

INTR 38866 ... 38663 OCKHAMG-CDNA

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 167

- Ceres seq\_id 1376587

- Alternative transcription start site(s) located in SEQ

ID NO 167:

8,9,10,12,13,14,15,16,17,30,34,39,41

30

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 168

- Ceres seq\_id 1376588

- Location of start within SEQ ID NO 167: at 2 nt.

35

## (C) Nomination and Annotation of Domains within

(Dp) Related Amino Acid Sequences

- Alignment No. 108

- gi No. 4539292

- Description:

- % Identity: 99.4

- Alignment Length: 177

- Location of Alignment in SEQ ID NO 168: from 26

to 202

45

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 169

- Ceres seq\_id 1376589

- Location of start within SEQ ID NO 167: at 77 nt.

50

163

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 109

- gi No. 4539292

- Description:

- % Identity: 99.4

- Alignment Length: 177

- Location of Alignment in SEQ ID NO 169: from 1 to

10 177

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 170

- Ceres seq\_id 1376590

- Location of start within SEQ ID NO 167: at 212 nt.

15

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 110

- gi No. 4539292

- Description:

- % Identity: 99.4

- Alignment Length: 177

- Location of Alignment in SEQ ID NO 170: from 1 to

25 132

(Ba) Polypeptide Activities: Similar 40S ribosomal protein activities.

30

Maximum Length Sequence:

related to:

Clone IDs:

112110

35 Public Genomic DNA:

gi No: 4263774

Predicted Exons:

INIT 1816 ... 1814 OCKHAMG-CDS

INTR 1290 ... 1112 OCKHAMG-CDS

TERM 959 ... 803 OCKHAMG-CDS

gi No: 4510360

Predicted Exons:

INIT 114707... 114705 OCKHAMG-CDS

INTR 114181... 114003 OCKHAMG-CDS

TERM 113850... 113694 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 171

- Ceres seq\_id 1378581

- Alternative transcription start site(s) located in SEQ

50 ID NO 171:

-35,-4,-3,18,20

164

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 172  
 - Ceres seq\_id 1378582  
 - Location of start within SEQ ID NO 171: at 112 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 172: at 17 aa.

- (C) Nomination and Annotation of Domains within

- Predicted Polypeptide(s)  
 - DnaJ domain  
 - Location within SEQ ID NO 172: from 57 to 101 aa.

- (Dp) Related Amino Acid Sequences

- Alignment No. 111  
 - gi No. 4263775  
 - Description:  
 - % Identity: 100  
 - Alignment Length: 112  
 - Location of Alignment in SEQ ID NO 172: from 1 to

112

- (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 173  
 - Ceres seq\_id 1378583  
 - Location of start within SEQ ID NO 171: at 256 nt.

- (C) Nomination and Annotation of Domains within

- Predicted Polypeptide(s)  
 - DnaJ domain  
 - Location within SEQ ID NO 173: from 9 to 53 aa.

- (Dp) Related Amino Acid Sequences

- Alignment No. 112  
 - gi No. 4263775  
 - Description:  
 - % Identity: 100  
 - Alignment Length: 112  
 - Location of Alignment in SEQ ID NO 173: from 1 to

64

- Maximum Length Sequence:

related to:

Clone IDs:

13599

- (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 174  
 - Ceres seq\_id 1383462  
 - Alternative transcription start site(s) located in SEQ ID NO 174:  
 -2,2,3,4,5,6,8,14,18,24

50

165

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 175  
 - Ceres seq\_id 1383463  
 - Location of start within SEQ ID NO 174: at 3 nt.

- (C) Nomination and Annotation of Domains within

- Predicted Polypeptide(s)  
 - Plant lipid transfer protein family  
 - Location within SEQ ID NO 175: from 46 to 119 aa.

- (Dp) Related Amino Acid Sequences

- Alignment No. 113  
 - gi No. 3128176  
 - Description:  
 - % Identity: 37  
 - Alignment Length: 194  
 - Location of Alignment in SEQ ID NO 175: from 23

to 197

20

- (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 176  
 - Ceres seq\_id 1383464  
 - Location of start within SEQ ID NO 174: at 42 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 176: at 24 aa.

- (C) Nomination and Annotation of Domains within

- Predicted Polypeptide(s)  
 - Plant lipid transfer protein family  
 - Location within SEQ ID NO 176: from 33 to 106 aa.

- (Dp) Related Amino Acid Sequences

- Alignment No. 114  
 - gi No. 3128176  
 - Description:  
 - % Identity: 37  
 - Alignment Length: 194  
 - Location of Alignment in SEQ ID NO 176: from 10

40 to 184

- (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 177  
 - Ceres seq\_id 1383465  
 - Location of start within SEQ ID NO 174: at 90 nt.

- (C) Nomination and Annotation of Domains within

- Predicted Polypeptide(s)  
 - Plant lipid transfer protein family  
 - Location within SEQ ID NO 177: from 17 to 90 aa.

50

- 166
- (Dp) Related Amino Acid Sequences
- Alignment No. 115
  - gi No. 3128176
  - Description: 37
  - % Identity: 37
  - Alignment Length: 194
  - Location of Alignment in SEQ ID NO 177: from 1 to 168
- 5
- 10 Maximum Length Sequence:  
related to:  
Clone IDs:  
156375
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 178
  - Ceres seq\_id 1386215
  - Alternative transcription start site(s) located in SEQ ID NO 178:
- 15
- 20 -38,12,17,18,19,20,26
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 179
  - Ceres seq\_id 1386216
  - Location of start within SEQ ID NO 178: at 2 nt.
- 25
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)
- Photosystem I psaG / psaK
  - Location within SEQ ID NO 179: from 77 to 138 aa.
- 30
- (Dp) Related Amino Acid Sequences
- Alignment No. 116
  - gi No. 3885511
  - Description: 79.3
  - % Identity: 79.3
  - Alignment Length: 112
  - Location of Alignment in SEQ ID NO 179: from 28 to 138
- 35
- 40 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 180
  - Ceres seq\_id 1386217
  - Location of start within SEQ ID NO 178: at 83 nt.
- 45
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)
- Photosystem I psaG / psaK
  - Location within SEQ ID NO 180: from 50 to 111 aa.
- 50
- (Dp) Related Amino Acid Sequences
- Alignment No. 117

- 167
- gi No. 3885511
  - Description: 79.3
  - Alignment Length: 112
  - Location of Alignment in SEQ ID NO 180: from 1 to 111
- 5
- 10 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 181
  - Ceres seq\_id 1386218
  - Location of start within SEQ ID NO 178: at 95 nt.
- 15
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)
- Photosystem I psaG / psaK
  - Location within SEQ ID NO 181: from 46 to 107 aa.
- 20
- (Dp) Related Amino Acid Sequences
- Alignment No. 118
  - gi No. 3885511
  - Description: 79.3
  - % Identity: 79.3
  - Alignment Length: 112
  - Location of Alignment in SEQ ID NO 181: from 1 to 107
- 25
- 30 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 182
  - Ceres seq\_id 2025156
  - Location of start within SEQ ID NO 178: at 502 nt.
- 35
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Maximum Length Sequence:  
related to:  
Clone IDs:  
21233
- 40 (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 183
  - Ceres seq\_id 1388499
  - Alternative transcription start site(s) located in SEQ ID NO 183:  
2,5,6,10,16,30,89,346,349
- 45
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 184
  - Ceres seq\_id 1388500
  - Location of start within SEQ ID NO 183: at 62 nt.
- 50



168

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 119
- gi No. 2829899
- Description:
- % Identity: 49.3
- Alignment Length: 150
- Location of Alignment in SEQ ID NO 184: from 2 to

5

10 150

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 185
- Ceres seq\_id 1388501
- Location of start within SEQ ID NO 183: at 122 nt.

15

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 120
- gi No. 2829899
- Description:
- % Identity: 49.3
- Alignment Length: 150
- Location of Alignment in SEQ ID NO 185: from 1 to

25 130

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 186
- Ceres seq\_id 1388502
- Location of start within SEQ ID NO 183: at 266 nt.

30

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 121
- gi No. 2829899
- Description:
- % Identity: 49.3
- Alignment Length: 150
- Location of Alignment in SEQ ID NO 186: from 1 to

40

82

(Ba) Polypeptide Activities: Similar to major latex protein activities.

45

Maximum Length Sequence:

related to:

Clone IDs:

50

169

21304

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 187
- Ceres seq\_id 1388519

5

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 188
- Ceres seq\_id 1388520
- Location of start within SEQ ID NO 187: at 1 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 188: at 38 aa.

10

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 122
- gi No. 4584110
- Description:
- % Identity: 48.5
- Alignment Length: 163
- Location of Alignment in SEQ ID NO 188: from 20

20

to 182

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 189
- Ceres seq\_id 1388521
- Location of start within SEQ ID NO 187: at 58 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 189: at 19 aa.

30

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 123
- gi No. 4584110
- Description:
- % Identity: 48.5
- Alignment Length: 163
- Location of Alignment in SEQ ID NO 189: from 1 to

40

163.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 190
- Ceres seq\_id 1388522
- Location of start within SEQ ID NO 187: at 73 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 190: at 14 aa.

45

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

50

170

## (Dp) Related Amino Acid Sequences

- Alignment No. 124
- gi No. 4584110
- Description:
- % Identity: 48.5
- Alignment Length: 163
- Location of Alignment in SEQ ID NO 190: from 1 to

5

158

- 10 (Ba) Polypeptide Activities: Similar to pollen specific protein activities.

## Maximum Length Sequence:

related to:

15 Clone IDs:

2153

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 191
- Ceres seq\_id 1388563
- Alternative transcription start site(s) located in SEQ ID NO 191: -3,-2,-1,13,15,146

20

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 192
- Ceres seq\_id 1388564
- Location of start within SEQ ID NO 191: at 2 nt.

25

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 193
- Ceres seq\_id 1388565
- Location of start within SEQ ID NO 191: at 306 nt.

35

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 125
- gi No. 3927834
- Description:
- % Identity: 31.5
- Alignment Length: 108
- Location of Alignment in SEQ ID NO 193: from 1 to

45

43

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 194
- Ceres seq\_id 1388566

50

171

- Location of start within SEQ ID NO 191: at 455 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 194: at 36 aa.

- 5 (Ba) Polypeptide Activities: Similar to hydroxproline-rich glycoprotein activities.

## Maximum Length Sequence:

10 related to:

Clone IDs:

22488

## Public Genomic DNA:

gi No: 5708384

## Predicted Exons:

- INIT 82228 ... 82063 OCKHAMG-CDS
- TERM 81730 ... 81381 OCKHAMG-CDS

15

gi No: 5732090

## Predicted Exons:

- INIT 21301 ... 21136 OCKHAMG-CDS
- TERM 20803 ... 20454 OCKHAMG-CDS

20

gi No: 5870169

## Predicted Exons:

- INIT 89258 ... 89423 OCKHAMG-CDS
- TERM 89756 ... 90105 OCKHAMG-CDS

25

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 195
- Ceres seq\_id 1388793
- Alternative transcription start site(s) located in SEQ ID NO 195:

30

7,14,15,20,24,30,33,42,44,58,63,64,66,76,94,97,98,99,101,102  
103,104,105,106,108,109,110,115,116,117

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 196
- Ceres seq\_id 1388794
- Location of start within SEQ ID NO 195: at 154 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 196: at 19 aa.

40

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 126
- gi No. 4584110
- Description:
- % Identity: 48.5
- Alignment Length: 163
- Location of Alignment in SEQ ID NO 196: from 1 to

45

50

163

172

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 197
- Ceres seq\_id 1388795
- Location of start within SEQ ID NO 195: at 169 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 197: at 14 aa.

5

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 127
- gi No. 4584110
- Description:
- % Identity: 48.5
- Alignment Length: 163
- Location of Alignment in SEQ ID NO 197: from 1 to

15

158

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 198
- Ceres seq\_id 1388796
- Location of start within SEQ ID NO 195: at 205 nt.

20

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 128
- gi No. 4584110
- Description:
- % Identity: 48.5
- Alignment Length: 163
- Location of Alignment in SEQ ID NO 198: from 1 to

30

146

## (Ba) Polypeptide Activities: Similar to pollen specific protein activities.

35

## Maximum Length Sequence:

related to:

Clone IDs:

26569

## Public Genomic DNA:

gi No: 6143856

## Predicted Exons:

- INTR 78628 ... 78255 OCKHAMG-CDNA
- INTR 78160 ... 78004 OCKHAMG-CDNA

45

- INTR 78624 ... 78255 OCKHAMG-CDNA
- INTR 78160 ... 78004 OCKHAMG-CDNA

50

173

INTR 78624 ... 78255 OCKHAMG-CDNA

INTR 78160 ... 78006 OCKHAMG-CDNA

INTR 78624 ... 78283 OCKHAMG-CDNA

gi No: 6223633

## Predicted Exons:

- INTR 78629 ... 78256 OCKHAMG-CDNA
- INTR 78161 ... 78005 OCKHAMG-CDNA

INTR 78625 ... 78256 OCKHAMG-CDNA

INTR 78161 ... 78005 OCKHAMG-CDNA

INTR 78625 ... 78256 OCKHAMG-CDNA

INTR 78161 ... 78007 OCKHAMG-CDNA

INTR 78625 ... 78284 OCKHAMG-CDNA

INTR 78587 ... 78256 OCKHAMG-CDS

TERM 78161 ... 78152 OCKHAMG-CDS

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 199
- Ceres seq\_id 1392041
- Alternative transcription start site(s) located in SEQ ID NO 199:

25 3,4,5,6,7,12

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 200
- Ceres seq\_id 1392042
- Location of start within SEQ ID NO 199: at 45 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 200: at 19 aa.

30

ID NO 200: at 19 aa.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Plant lipid transfer protein family
- Location within SEQ ID NO 200: from 23 to 112 aa.

35

## (Dp) Related Amino Acid Sequences

- Alignment No. 129
- gi No. 2497753
- Description:
- % Identity: 47
- Alignment Length: 117
- Location of Alignment in SEQ ID NO 200: from 4 to

45

113

## Maximum Length Sequence:

related to:

Clone IDs:

28475

50

174

## Public Genomic DNA:

gi No: 5731404

## Predicted Exons:

5 INTR 70440 ... 70346 OCKHAMG-CDNA  
 INTR 70076 ... 69865 OCKHAMG-CDNA  
 INTR 69508 ... 69434 OCKHAMG-CDNA  
 INTR 69358 ... 69284 OCKHAMG-CDNA  
 INTR 69195 ... 69115 OCKHAMG-CDNA  
 INTR 69033 ... 68959 OCKHAMG-CDNA  
 INTR 68868 ... 68695 OCKHAMG-CDNA  
 INTR 68581 ... 68344 OCKHAMG-CDNA

15 INTR 70056 ... 69865 OCKHAMG-CDS  
 INTR 69508 ... 69434 OCKHAMG-CDS  
 INTR 69358 ... 69284 OCKHAMG-CDS  
 INTR 69195 ... 69115 OCKHAMG-CDS  
 INTR 69033 ... 68959 OCKHAMG-CDS  
 INTR 68868 ... 68695 OCKHAMG-CDS  
 TERM 68581 ... 68546 OCKHAMG-CDS

## 20 (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 201

- Ceres seq\_id 1393342

- Alternative transcription start site(s) located in SEQ

ID NO 201:

16,48

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 202

- Ceres seq\_id 1393343

- Location of start within SEQ ID NO 201: at 118 nt.

(C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 130

- gi No. 3386621

- Description:

- % Identity: 83

- Alignment Length: 224

- Location of Alignment in SEQ ID NO 202: from 2 to

225

(Ba) Polypeptide Activities: Arabidopsis specific gene,

dicot specific gene, plant specific gene.

## Maximum Length Sequence:

related to:

Clone IDs:

29453

## 50 (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 203

175

- Ceres seq\_id 1393556

- Alternative transcription start site(s) located in SEQ

ID NO 203:

7,11,14,23,37,41,42,57,61

5

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 204

- Ceres seq\_id 1393557

- Location of start within SEQ ID NO 203: at 3 nt.

10

(C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 131

- gi No. 1082054

- Description:

- % Identity: 36.1

- Alignment Length: 61

- Location of Alignment in SEQ ID NO 204: from 23

20 to 82

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 205

- Ceres seq\_id 1393558

- Location of start within SEQ ID NO 203: at 69 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 205: at 41 aa.

(C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 132

- gi No. 1082054

- Description:

- % Identity: 36.1

- Alignment Length: 61

- Location of Alignment in SEQ ID NO 205: from 1 to

60

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 206

- Ceres seq\_id 1393559

- Location of start within SEQ ID NO 203: at 72 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 206: at 40 aa.

(C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 133

- gi No. 1082054

- 176
- Description:
  - % Identity: 36.1
  - Alignment Length: 61
  - Location of Alignment in SEQ ID NO 206: from 1 to
- 5 59
- (Ba) Polypeptide Activities: Similar to transmembrane copper transporter protein activities.
- 10
- Maximum Length Sequence:  
related to:  
Clone IDs:  
42384
- 15 (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 207
  - Ceres seq\_id 1396782
  - Alternative transcription start site(s) located in SEQ ID NO 207:
- 20 -1,29
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 208
  - Ceres seq\_id 1396783
  - Location of start within SEQ ID NO 207: at 3 nt.
- 25
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences
- Alignment No. 134
  - gi No. 4512613
  - Description:
  - % Identity: 99.2
  - Alignment Length: 122
  - Location of Alignment in SEQ ID NO 208: from 11 to 132
- 35
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 209
  - Ceres seq\_id 1396784
  - Location of start within SEQ ID NO 207: at 33 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 209: at 19 aa.
- 40
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences
- Alignment No. 135
  - gi No. 4512613
  - Description:
  - % Identity: 99.2
- 45
- 50

- 177
- Alignment Length: 122
  - Location of Alignment in SEQ ID NO 209: from 1 to
- 122
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 210
  - Ceres seq\_id 1396785
  - Location of start within SEQ ID NO 207: at 45 nt.
- 5
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences
- Alignment No. 136
  - gi No. 4512613
  - Description:
  - % Identity: 99.2
  - Alignment Length: 122
  - Location of Alignment in SEQ ID NO 210: from 1 to
- 10
- 118
- (Ba) Polypeptide Activities: Similar to intergenic region of cyt2-MDH1 of yeast activities.
- 20
- Maximum Length Sequence:  
related to:  
Clone IDs:  
42402
- 25
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 211
  - Ceres seq\_id 1396802
  - Alternative transcription start site(s) located in SEQ ID NO 211:  
-145,-56
- 30
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 212
  - Ceres seq\_id 1396803
  - Location of start within SEQ ID NO 211: at 1 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 212: at 19 aa.
- 35
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences
- Alignment No. 137
  - gi No. 2463339
  - Description:
  - % Identity: 38.6
  - Alignment Length: 70
  - Location of Alignment in SEQ ID NO 212: from 136 to 205
- 40
- 45
- 50

178

- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 213  
- Ceres seq\_id 1396804  
- Location of start within SEQ ID NO 211: at 136 nt.

5

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 138
- gi No. 2463339
- Description:
- % Identity: 38.6
- Alignment Length: 70
- Location of Alignment in SEQ ID NO 213: from 91 to 160

15

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 214

- Ceres seq\_id 1396805

- Location of start within SEQ ID NO 211: at 172 nt.

20

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 139

- gi No. 2463339

- Description:

- % Identity: 38.6

- Alignment Length: 70

- Location of Alignment in SEQ ID NO 214: from 79 to 148

30

(Ba) Polypeptide Activities: Similar to rbcX protein activities.

35

Maximum Length Sequence:

related to:

Clone IDs:

5105

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 215

- Ceres seq\_id 1397130

- Alternative transcription start site(s) located in SEQ ID NO 215:

-1

45

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 216

- Ceres seq\_id 1397131

- Location of start within SEQ ID NO 215: at 1 nt.

50

179

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 140
- gi No. 2827551
- Description:
- % Identity: 63.1
- Alignment Length: 179
- Location of Alignment in SEQ ID NO 216: from 2 to 176

10

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 217

- Ceres seq\_id 1397132

- Location of start within SEQ ID NO 215: at 52 nt.

15

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 141

- gi No. 2827551

- Description:

- % Identity: 63.1

- Alignment Length: 179

- Location of Alignment in SEQ ID NO 217: from 1 to 159

25

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 218

- Ceres seq\_id 1397133

- Location of start within SEQ ID NO 215: at 211 nt.

30

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 142

- gi No. 2827551

- Description:

- % Identity: 63.1

- Alignment Length: 179

- Location of Alignment in SEQ ID NO 218: from 1 to 106

40

(Ba) Polypeptide Activities: Similar to meth Cpg binding protein activities.

45

Maximum Length Sequence:

related to:

Clone IDs:

50

180

92204

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 219
- Ceres seq\_id 1398004
- Alternative transcription start site(s) located in SEQ ID NO 219:

2,3,5,8,28

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 220
- Ceres seq\_id 1398005
- Location of start within SEQ ID NO 219: at 68 nt.

## (C) Nomination and Annotation of Domains within

## (Dp) Related Amino Acid Sequences

- Alignment No. 143
- gi No. 4454037
- Description:
- % Identity: 99.3
- Alignment Length: 151
- Location of Alignment in SEQ ID NO 220: from 1 to 151

- (Ba) Polypeptide Activities: Similar to major latex protein activities.

Maximum Length Sequence:  
related to:

Clone IDs:

98584

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 221
- Ceres seq\_id 1399370

35

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 222
- Ceres seq\_id 1399371
- Location of start within SEQ ID NO 221: at 3 nt.

## (C) Nomination and Annotation of Domains within

## (Dp) Related Amino Acid Sequences

- Alignment No. 144
- gi No. 4886285
- Description:
- % Identity: 28.8
- Alignment Length: 59
- Location of Alignment in SEQ ID NO 222: from 25

50 to 83

181

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 223
- Ceres seq\_id 1399372
- Location of start within SEQ ID NO 221: at 45 nt.

5

## (C) Nomination and Annotation of Domains within

## (Dp) Related Amino Acid Sequences

- Alignment No. 145
- gi No. 4886285
- Description:
- % Identity: 28.8
- Alignment Length: 59
- Location of Alignment in SEQ ID NO 223: from 11

15 to 69

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 224
- Ceres seq\_id 1399373
- Location of start within SEQ ID NO 221: at 60 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 224: at 25 aa.

20

## (C) Nomination and Annotation of Domains within

## (Dp) Related Amino Acid Sequences

- Alignment No. 146
- gi No. 4886285
- Description:
- % Identity: 28.8
- Alignment Length: 59
- Location of Alignment in SEQ ID NO 224: from 6 to

64

- (Ba) Polypeptide Activities: Similar to outer envelope membrane protein in chloroplast in pea activities.

Maximum Length Sequence:

related to:

Clone IDs:

107400

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 225
- Ceres seq\_id 1425147
- Alternative transcription start site(s) located in SEQ ID NO 225:

20,31,49,58,80

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 226

- 182
- Ceres seq\_id 1425148
- Location of start within SEQ ID NO 225: at 85 nt.
- (C) Nomination and Annotation of Domains within
- (Dp) Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 147
- gi No. 3510256
- Description:
- % Identity: 37.8
- Alignment Length: 158
- Location of Alignment in SEQ ID NO 226: from 1 to
- 155
- 15 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 227
- Ceres seq\_id 1425149
- Location of start within SEQ ID NO 225: at 241 nt.
- (C) Nomination and Annotation of Domains within
- (Dp) Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 148
- gi No. 3510256
- Description:
- % Identity: 37.8
- Alignment Length: 158
- Location of Alignment in SEQ ID NO 227: from 1 to
- 103
- 20 (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot specific gene, plant specific gene.
- Maximum Length Sequence:
- related to:
- Clone IDs:
- 11073
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 228
- Ceres seq\_id 1441102
- Alternative transcription start site(s) located in SEQ ID NO 228:
- 2,3,4,5,8,10,16,22,51,69,388
- 45 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 229
- Ceres seq\_id 1441103
- Location of start within SEQ ID NO 228: at 1 nt.
- (C) Nomination and Annotation of Domains within
- Predicted Polypeptide(s)
- 50

- 183
- (Dp) Related Amino Acid Sequences
- Alignment No. 149
- gi No. 4337175
- Description:
- % Identity: 100
- Alignment Length: 179
- Location of Alignment in SEQ ID NO 229: from 19 to 197
- 10 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 230
- Ceres seq\_id 1441104
- Location of start within SEQ ID NO 228: at 55 nt.
- (C) Nomination and Annotation of Domains within
- Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 150
- gi No. 4337175
- Description:
- % Identity: 100
- Alignment Length: 179
- Location of Alignment in SEQ ID NO 230: from 1 to
- 179
- 25 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 231
- Ceres seq\_id 1441105
- Location of start within SEQ ID NO 228: at 94 nt.
- (C) Nomination and Annotation of Domains within
- Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 151
- gi No. 4337175
- Description:
- % Identity: 100
- Alignment Length: 179
- Location of Alignment in SEQ ID NO 231: from 1 to
- 166
- 40 (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot specific gene, plant specific gene.
- Maximum Length Sequence:
- related to:
- Clone IDs:
- 267227
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 232
- Ceres seq\_id 1447480
- 50



184

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 233  
 - Ceres seq\_id 1447481  
 - Location of start within SEQ ID NO 232: at 52 nt.
- (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 152  
 - gi No. 3510256  
 - Description:  
 - % Identity: 54  
 - Alignment Length: 50  
 - Location of Alignment in SEQ ID NO 233: from 77  
 to 126
- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 234  
 - Ceres seq\_id 1447482  
 - Location of start within SEQ ID NO 232: at 208 nt.
- (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 153  
 - gi No. 3510256  
 - Description:  
 - % Identity: 54  
 - Alignment Length: 50  
 - Location of Alignment in SEQ ID NO 234: from 25  
 to 74

(Ba) Polypeptide Activities: Arabidopsis specific gene,  
 dicot specific gene, plant specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

268712

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 235

- Ceres seq\_id 1447577

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 236

- Ceres seq\_id 1447578

- Location of start within SEQ ID NO 235: at 2 nt.

185

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 154

- gi No. 5080769

- Description:

- % Identity: 100

- Alignment Length: 17

- Location of Alignment in SEQ ID NO 236: from 16  
 to 32

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 237

- Ceres seq\_id 1447579

- Location of start within SEQ ID NO 235: at 165 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 155

- gi No. 5080769

- Description:

- % Identity: 91.5

- Alignment Length: 153

- Location of Alignment in SEQ ID NO 237: from 1 to  
 115

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 238

- Ceres seq\_id 1447580

- Location of start within SEQ ID NO 235: at 298 nt.

(Ba) Polypeptide Activities: Similar to  
 hydroxyproline-rich glycoprotein activities.

Maximum Length Sequence:

related to:

Clone IDs:

94821

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 239

- Ceres seq\_id 1447922

- Alternative transcription start site(s) located in SEQ  
 ID NO 239:  
 12,16

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 240

- Ceres seq\_id 1447923

- Location of start within SEQ ID NO 239: at 81 nt.

186

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 5  
- Alignment No. 156  
- Description:  
- % Identity: 37.8  
- Alignment Length: 158  
- Location of Alignment in SEQ ID NO 240: from 1 to 155
- 10  
- Location of Alignment in SEQ ID NO 240: from 1 to 155
- 15  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 241  
- Ceres seq\_id 1447924  
- Location of start within SEQ ID NO 239: at 237 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 20  
- Alignment No. 157  
- Description:  
- % Identity: 37.8  
- Alignment Length: 158  
- Location of Alignment in SEQ ID NO 241: from 1 to 103
- 25  
- Location of Alignment in SEQ ID NO 241: from 1 to 103

- 30 (Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot specific gene, plant specific gene.

## Maximum Length Sequence:

related to:

Clone IDs:

20539

## (Ac) cDNA Polynucleotide Sequence

- 35  
- Pat. Appln. SEQ ID NO 242  
- Ceres seq\_id 1448012  
- Alternative transcription start site(s) located in SEQ ID NO 242:  
-30,-6

## (B) Polypeptide Sequence

- 45  
- Pat. Appln. SEQ ID NO 243  
- Ceres seq\_id 1448013  
- Location of start within SEQ ID NO 242: at 79 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- 50  
- Ribosomal protein S16  
- Location within SEQ ID NO 243: from 9 to 70 aa.

187

## (Dp) Related Amino Acid Sequences

## - Alignment No. 158

- 5  
- gi No. 3096931  
- Description:  
- % Identity: 57.5  
- Alignment Length: 113  
- Location of Alignment in SEQ ID NO 243: from 1 to 112
- 10  
- Location of Alignment in SEQ ID NO 243: from 1 to 112

## (B) Polypeptide Sequence

## - Pat. Appln. SEQ ID NO 244

- 15  
- Ceres seq\_id 1448014  
- Location of start within SEQ ID NO 242: at 139 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## - Ribosomal protein S16

- 20  
- Location within SEQ ID NO 244: from 1 to 50 aa.

## (Dp) Related Amino Acid Sequences

## - Alignment No. 159

- 25  
- gi No. 3096931  
- Description:  
- % Identity: 57.5  
- Alignment Length: 113  
- Location of Alignment in SEQ ID NO 244: from 1 to 92
- 92  
- Location of Alignment in SEQ ID NO 244: from 1 to 92

## (B) Polypeptide Sequence

## - Pat. Appln. SEQ ID NO 245

- 30  
- Ceres seq\_id 1448015  
- Location of start within SEQ ID NO 242: at 232 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

## - Alignment No. 160

- 35  
- gi No. 3096931  
- Description:  
- % Identity: 57.5  
- Alignment Length: 113  
- Location of Alignment in SEQ ID NO 245: from 1 to 61
- 61  
- Location of Alignment in SEQ ID NO 245: from 1 to 61

## Maximum Length Sequence:

related to:

Clone IDs:

34091

## (Ac) cDNA Polynucleotide Sequence

- 50  
- Pat. Appln. SEQ ID NO 246

188

- Ceres seq\_id 1448135
- Alternative transcription start site(s) located in SEQ ID NO 246:

-1

5

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 247
- Ceres seq\_id 1448136
- Location of start within SEQ ID NO 246: at 3 nt.

10

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

15

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 248
- Ceres seq\_id 1448137
- Location of start within SEQ ID NO 246: at 68 nt.

20

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein S14p/S29e
- Location within SEQ ID NO 248: from 3 to 54 aa.

25

## (Dp) Related Amino Acid Sequences

- Alignment No. 161
- gi No. 4506717
- Description:
- % Identity: 72.2
- Alignment Length: 54
- Location of Alignment in SEQ ID NO 248: from 1 to

54

35

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 249
- Ceres seq\_id 1448138
- Location of start within SEQ ID NO 246: at 85 nt.

40

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

## Maximum Length Sequence:

related to:

Clone IDs:

39285

Public Genomic DNA:

gi No: 4406776

Predicted Exons:

SINGLE 61584 ... 61895 GENBANK

50

189

SINGLE 61584 ... 61895 OCKHAMG-CDS

- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 250
- Ceres seq\_id 1448185
- Alternative transcription start site(s) located in SEQ ID NO 250:

-36,-18,-17,16,19

10

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 251
- Ceres seq\_id 1448186
- Location of start within SEQ ID NO 250: at 41 nt.

15

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 162
- gi No. 4406787
- Description:
- % Identity: 100
- Alignment Length: 103
- Location of Alignment in SEQ ID NO 251: from 1 to

103

25

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 252
- Ceres seq\_id 1448187
- Location of start within SEQ ID NO 250: at 68 nt.

30

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 163
- gi No. 4406787
- Description:
- % Identity: 100
- Alignment Length: 103
- Location of Alignment in SEQ ID NO 252: from 1 to

94

40

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 253
- Ceres seq\_id 1448188
- Location of start within SEQ ID NO 250: at 89 nt.

45

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 164
- gi No. 4406787
- Description:

50

- 190
- % Identity: 100
  - Alignment Length: 103
  - Location of Alignment in SEQ ID NO 253: from 1 to 87
- 5 (Ba) Polypeptide Activities: Similar to NADH dehydrogenase protein activities
- 10 Maximum Length Sequence:  
related to:  
Clone IDs: 99119
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 254
  - Ceres seq\_id 1450875
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 255
  - Ceres seq\_id 1450876
  - Location of start within SEQ ID NO 254: at 2 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 255: at 26 aa.
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- ATPases associated with various cellular activities (AAA)
  - Location within SEQ ID NO 255: from 391 to 598
- aa.
- (Dp) Related Amino Acid Sequences
- Alignment No. 165
  - gi No. 4309734
  - Description:
  - % Identity: 86
  - Alignment Length: 633
  - Location of Alignment in SEQ ID NO 255: from 9 to 634
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 256
  - Ceres seq\_id 1450877
  - Location of start within SEQ ID NO 254: at 14 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 256: at 22 aa.
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- 191
- ATPases associated with various cellular activities (AAA)
  - Location within SEQ ID NO 256: from 387 to 594
- aa.
- (Dp) Related Amino Acid Sequences
- Alignment No. 166
  - gi No. 4309734
  - Description:
  - % Identity: 86
  - Alignment Length: 633
  - Location of Alignment in SEQ ID NO 256: from 5 to 630
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 257
  - Ceres seq\_id 1450878
  - Location of start within SEQ ID NO 254: at 68 nt.
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- ATPases associated with various cellular activities (AAA)
  - Location within SEQ ID NO 257: from 369 to 576
- aa.
- (Dp) Related Amino Acid Sequences
- Alignment No. 167
  - gi No. 4309734
  - Description:
  - % Identity: 86
  - Alignment Length: 633
  - Location of Alignment in SEQ ID NO 257: from 1 to 612
- Maximum Length Sequence:  
related to:  
Clone IDs: 269321
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 258
  - Ceres seq\_id 1459191
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 259
  - Ceres seq\_id 1459192
  - Location of start within SEQ ID NO 258: at 1 nt.
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

192

## (Dp) Related Amino Acid Sequences

- Alignment No. 168
- gi No. 4490728
- Description:
- % Identity: 100
- Alignment Length: 17
- Location of Alignment in SEQ ID NO 259: from 21 to 37

5

## 10 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 260
- Ceres seq\_id 1459193
- Location of start within SEQ ID NO 258: at 3 nt.

## 15 (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 169
- gi No. 4490728
- Description:
- % Identity: 85.9
- Alignment Length: 85
- Location of Alignment in SEQ ID NO 260: from 37 to 110

25

## (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot specific gene, plant specific gene.

## Maximum Length Sequence:

related to:

Clone IDs:

8446

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 261
- Ceres seq\_id 1461848
- Alternative transcription start site(s) located in SEQ ID NO 261:  
-3,6,7,11,31,32,34,62,63,69,70,71,97

40

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 262
- Ceres seq\_id 1461849
- Location of start within SEQ ID NO 261: at 113 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 262: at 36 aa.

45

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- MAPEG family
- Location within SEQ ID NO 262: from 59 to 136 aa.

50

193

## (Dp) Related Amino Acid Sequences

- Alignment No. 170
- gi No. 4758714
- Description:
- % Identity: 41.9
- Alignment Length: 131
- Location of Alignment in SEQ ID NO 262: from 8 to 136

5

## 10 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 263
- Ceres seq\_id 1461850
- Location of start within SEQ ID NO 261: at 197 nt.

## 15 (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- MAPEG family
- Location within SEQ ID NO 263: from 31 to 108 aa.

20

## (Dp) Related Amino Acid Sequences

- Alignment No. 171
- gi No. 4758714
- Description:
- % Identity: 41.9
- Alignment Length: 131
- Location of Alignment in SEQ ID NO 263: from 1 to 108

25

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 264
- Ceres seq\_id 1461851
- Location of start within SEQ ID NO 261: at 329 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 264: at 22 aa.

35

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- MAPEG family
- Location within SEQ ID NO 264: from 1 to 64 aa.

40

## (Dp) Related Amino Acid Sequences

- Alignment No. 172
- gi No. 4758714
- Description:
- % Identity: 41.9
- Alignment Length: 131
- Location of Alignment in SEQ ID NO 264: from 1 to 64

45

50 Maximum Length Sequence:  
related to:

194

## Clone IDs:

25093

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 265

- Ceres seq\_id 1472772

- Alternative transcription start site(s) located in SEQ

ID NO 265:

-1,3

10 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 266

- Ceres seq\_id 1472773

- Location of start within SEQ ID NO 265: at 3 nt.

- Location of Signal Peptide Cleavage Site within SEQ

15 ID NO 266: at 32 aa.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Plant lipid transfer protein family

20 - Location within SEQ ID NO 266: from 45 to 108 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 173

- gi No. 3062791

- Description:

- % Identity: 72.2

- Alignment Length: 90

- Location of Alignment in SEQ ID NO 266: from 21

to 110

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 267

- Ceres seq\_id 1472774

- Location of start within SEQ ID NO 265: at 27 nt.

- Location of Signal Peptide Cleavage Site within SEQ

35 ID NO 267: at 24 aa.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Plant lipid transfer protein family

40 - Location within SEQ ID NO 267: from 37 to 100 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 174

- gi No. 3062791

- Description:

- % Identity: 72.2

- Alignment Length: 90

- Location of Alignment in SEQ ID NO 267: from 13

50 to 102

SUBSTITUTE SHEET (RULE 26)

195

## Maximum Length Sequence:

related to:

Clone IDs:

42300

5 (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 268

- Ceres seq\_id 1533352

- Alternative transcription start site(s) located in SEQ

ID NO 268:

10 17

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 269

- Ceres seq\_id 1533353

15 - Location of start within SEQ ID NO 268: at 1 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- haloacid dehalogenase-like hydrolase

20 - Location within SEQ ID NO 269: from 82 to 266 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 175

- gi No. 3913203

- Description:

- % Identity: 32.8

- Alignment Length: 196

- Location of Alignment in SEQ ID NO 269: from 81

to 266

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 270

- Ceres seq\_id 1533354

35 - Location of start within SEQ ID NO 268: at 25 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- haloacid dehalogenase-like hydrolase

40 - Location within SEQ ID NO 270: from 74 to 258 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 176

- gi No. 3913203

- Description:

- % Identity: 32.8

- Alignment Length: 196

- Location of Alignment in SEQ ID NO 270: from 73

to 258

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 271

SUBSTITUTE SHEET (RULE 26)

196

- Ceres seq\_id 1533355
  - Location of start within SEQ ID NO 268: at 265 nt.
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)
- haloacid dehalogenase-like hydrolase
  - Location within SEQ ID NO 271: from 1 to 178 aa.

5

## (Dp) Related Amino Acid Sequences

- Alignment No. 177
- gi No. 3913203
- Description:
- % Identity: 32.8
- Alignment Length: 196
- Location of Alignment in SEQ ID NO 271: from 1 to 178

15

## Maximum Length Sequence:

related to:

Clone IDs:

158412

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 272
- Ceres seq\_id 1534544

25

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 273
- Ceres seq\_id 1534545
- Location of start within SEQ ID NO 272: at 2 nt.

30

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 178
- gi No. 2317676
- Description:
- % Identity: 37.2
- Alignment Length: 115
- Location of Alignment in SEQ ID NO 273: from 22

40

to 134

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 274
- Ceres seq\_id 1534546
- Location of start within SEQ ID NO 272: at 524 nt.

45

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 179

50

197

- gi No. 4836939
- Description:
- % Identity: 47.3
- Alignment Length: 150
- Location of Alignment in SEQ ID NO 274: from 19

5

to 166

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 275
- Ceres seq\_id 1534547
- Location of start within SEQ ID NO 272: at 815. nt.

10

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 180
- gi No. 4836939
- Description:
- % Identity: 47.3
- Alignment Length: 150
- Location of Alignment in SEQ ID NO 275: from 1 to

20

69

(Ba) Polypeptide Activities: Similar to calcium independent phospholipase A2 protein activities, and acy-protein thioesterase activities.

25

## Maximum Length Sequence:

related to:

Clone IDs:

108109

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 276
- Ceres seq\_id 1567172
- Alternative transcription start site(s) located in SEQ ID NO 276:

35

22

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 277
- Ceres seq\_id 1567173
- Location of start within SEQ ID NO 276: at 152 nt.

40

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 181
- gi No. 1076301

45

198  
 - Description: CONSTANS protein - Arabidopsis  
 thaliana >gi|1161514|emb|CAA64407| (X94937) CONSTANS protein  
 [Arabidopsis thaliana]  
 - % Identity: 49.2  
 - Alignment Length: 63  
 - Location of Alignment in SEQ ID NO 277: from 12  
 to 74

5

(B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 278  
 - Ceres seq\_id 1567174  
 - Location of start within SEQ ID NO 276: at 155 nt.

10

(C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 182  
 - gi No. 1076301  
 - Description: CONSTANS protein - Arabidopsis  
 thaliana >gi|1161514|emb|CAA64407| (X94937) CONSTANS protein  
 [Arabidopsis thaliana]  
 - % Identity: 49.2  
 - Alignment Length: 63  
 - Location of Alignment in SEQ ID NO 278: from 11  
 to 73

25

(B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 279  
 - Ceres seq\_id 1567175  
 - Location of start within SEQ ID NO 276: at 443 nt.

30

(C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 183  
 - gi No. 1076301  
 - Description: CONSTANS protein - Arabidopsis  
 thaliana >gi|1161514|emb|CAA64407| (X94937) CONSTANS protein  
 [Arabidopsis thaliana]  
 - % Identity: 67.4  
 - Alignment Length: 43  
 - Location of Alignment in SEQ ID NO 279: from 261  
 to 303

40

Maximum Length Sequence:  
 related to:  
 Clone IDs:  
 168  
 Public Genomic DNA:  
 gi No: 4895213  
 Predicted Exons:

50

199  
 INIT 22484 ... 22368 OCKHAMG-CDS  
 TERM 21760 ... 21704 OCKHAMG-CDS  
 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 280  
 - Ceres seq\_id 1567535  
 - Alternative transcription start site(s) located in SEQ  
 ID NO 280:  
 12,22,23,25,28,36,68

5

(B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 281  
 - Ceres seq\_id 1567536  
 - Location of start within SEQ ID NO 280: at 104 nt.

10

(C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 184  
 - gi No. 1351365  
 - Description: UBIQUINOL-CYTOCHROME C REDUCTASE  
 COMPLEX 6.7 KD PROTEIN (CR6) >gi|2130002|pir|IS68969  
 ubiquinol--cytochrome-c reductase (EC 1.10.2.2) - potato  
 >gi|633683|emb|CAA57768| (X82325) cytochrome c reductase  
 subunit [Solanum tuberosum]  
 - % Identity: 52.6  
 - Alignment Length: 58  
 - Location of Alignment in SEQ ID NO 281: from 2 to  
 57

25

Maximum Length Sequence:  
 related to:  
 Clone IDs:  
 21305  
 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 282  
 - Ceres seq\_id 1569689  
 - Alternative transcription start site(s) located in SEQ  
 ID NO 282:  
 28,29,49

35

(B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 283  
 - Ceres seq\_id 1569690  
 - Location of start within SEQ ID NO 282: at 1 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ  
 ID NO 283: at 32 aa.

40

(C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences

50



: 200

201

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 284  
 - Ceres seq\_id 1569691  
 - Location of start within SEQ ID NO 282: at 92 nt.
- (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 185  
 - gi No. 2191138  
 - Description: (AF007269) A\_IG002N01.18 gene  
 product [Arabidopsis thaliana]  
 - % Identity: 19.5  
 - Alignment Length: 87  
 - Location of Alignment in SEQ ID NO 284: from 3 to
- Maximum Length Sequence:  
 related to:  
 Clone IDs:  
 16143  
 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 285  
 - Ceres seq\_id 1569692  
 - Location of start within SEQ ID NO 282: at 248 nt.
- (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 186  
 - gi No. 2191138  
 - Description: (AF007269) A\_IG002N01.18 gene  
 product [Arabidopsis thaliana]  
 - % Identity: 19.5  
 - Alignment Length: 87  
 - Location of Alignment in SEQ ID NO 285: from 1 to
- Maximum Length Sequence:  
 related to:  
 Clone IDs:  
 41988  
 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 286  
 - Ceres seq\_id 1571042  
 - Alternative transcription start site(s) located in SEQ ID NO 286:  
 -1,2,3,4,7
- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 287  
 - Ceres seq\_id 1571043
- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 289  
 - Ceres seq\_id 1571080  
 - Location of start within SEQ ID NO 288: at 51 nt.
- (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 - Pathogenesis-related protein Bet v I family  
 - Location within SEQ ID NO 289: from 5 to 155 aa.
- (Dp) Related Amino Acid Sequences  
 - Alignment No. 188  
 - gi No. 1321731  
 - Description: (Z72439) major allergen Cor a 1  
 [Corylus avellana]  
 - % Identity: 35.7  
 - Alignment Length: 159  
 - Location of Alignment in SEQ ID NO 289: from 5 to
- Maximum Length Sequence:  
 related to:  
 Clone IDs:  
 42101  
 (Ac) cDNA Polynucleotide Sequence

202

- Pat. Appln. SEQ ID NO 290
- Ceres seq\_id 1572097

5

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 291
  - Ceres seq\_id 1572098
  - Location of start within SEQ ID NO 290: at 63 nt.

10

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- Adhesion lipoprotein
  - Location within SEQ ID NO 291: from 118 to 191

15

- (Dp) Related Amino Acid Sequences
- Alignment No. 189
  - gi No. 2129641

20

- Arabidopsis thaliana >gi1107493|emb|CAA63026| (X91960) major latex protein type1 [Arabidopsis thaliana]
- % Identity: 69.5
  - Alignment Length: 154
  - Location of Alignment in SEQ ID NO 291: from 1 to 154

25

- (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 292
- Ceres seq\_id 1572099
- Location of start within SEQ ID NO 290: at 336 nt.

30

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Adhesion lipoprotein
- Location within SEQ ID NO 292: from 27 to 100 aa.

35

- (Dp) Related Amino Acid Sequences

- Alignment No. 190
- gi No. 2129641
- Description: major latex protein type 1 - Arabidopsis thaliana >gi1107493|emb|CAA63026| (X91960) major latex protein type1 [Arabidopsis thaliana]
- % Identity: 69.5
- Alignment Length: 154
- Location of Alignment in SEQ ID NO 292: from 1 to 63

45

Maximum Length Sequence:  
related to:  
Clone IDs:  
27643

50

203

- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 293
  - Ceres seq\_id 1572890
  - Alternative transcription start site(s) located in SEQ ID NO 293:  
-10

5

- (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 294
- Ceres seq\_id 1572891
- Location of start within SEQ ID NO 293: at 1 nt.

10

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Photosystem I psag / psak
- Location within SEQ ID NO 294: from 71 to 151 aa.

15

- (Dp) Related Amino Acid Sequences

- Alignment No. 191
- gi No. 3885511
- Description: (AF084200) similar to PSI-K subunit of photosystem I from barley [Medicago sativa]
- % Identity: 80.3
- Alignment Length: 128
- Location of Alignment in SEQ ID NO 294: from 22 to 148

25

- (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 295
- Ceres seq\_id 1572892
- Location of start within SEQ ID NO 293: at 64 nt.

30

- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Photosystem I psag / psak
- Location within SEQ ID NO 295: from 50 to 130 aa.

35

- (Dp) Related Amino Acid Sequences

- Alignment No. 192
- gi No. 3885511
- Description: (AF084200) similar to PSI-K subunit of photosystem I from barley [Medicago sativa]
- % Identity: 80.3
- Alignment Length: 128
- Location of Alignment in SEQ ID NO 295: from 1 to 127

45

- (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 296
- Ceres seq\_id 1572893
- Location of start within SEQ ID NO 293: at 76 nt.

50

204

- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)
- Photosystem I psaG / psaK
  - Location within SEQ ID NO 296: from 46 to 126 aa.

5

- (Dp) Related Amino Acid Sequences
- Alignment No. 193
  - gi No. 3885511
  - Description: (AF084200) similar to PSI-K subunit of photosystem I from barley [Medicago sativa]
  - % Identity: 80.3
  - Alignment Length: 128
  - Location of Alignment in SEQ ID NO 296: from 1 to

15 123

Maximum Length Sequence:  
related to:

Clone IDs:

20 33027

- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 297
  - Ceres seq\_id 1573606
  - Alternative transcription start site(s) located in SEQ ID NO 297:  
7,8,13

25

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 298
- Ceres seq\_id 1573607
- Location of start within SEQ ID NO 297: at 56 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 298: at 19 aa.

30

(C) Nomination and Annotation of Domains within

- Predicted Polypeptide(s)
- Pollen proteins Ole e I family
  - Location within SEQ ID NO 298: from 32 to 131 aa.

35

(Dp) Related Amino Acid Sequences

- Alignment No. 194
- gi No. 4584110
- Description: (AJ133639) SAH7 protein [Arabidopsis thaliana]
- % Identity: 48.5
- Alignment Length: 163
- Location of Alignment in SEQ ID NO 298: from 1 to

45

163

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 299

50

205

- Ceres seq\_id 1573608
- Location of start within SEQ ID NO 297: at 71 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 299: at 14 aa.

5

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Pollen proteins Ole e I family
- Location within SEQ ID NO 299: from 27 to 126 aa.

10

(Dp) Related Amino Acid Sequences

- Alignment No. 195
- gi No. 4584110
- Description: (AJ133639) SAH7 protein [Arabidopsis thaliana]
- % Identity: 48.5
- Alignment Length: 163
- Location of Alignment in SEQ ID NO 299: from 1 to

15

158

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 300
- Ceres seq\_id 1573609
- Location of start within SEQ ID NO 297: at 107 nt.

20

25

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Pollen proteins Ole e I family
- Location within SEQ ID NO 300: from 15 to 114 aa.

30

(Dp) Related Amino Acid Sequences

- Alignment No. 196
- gi No. 4584110
- Description: (AJ133639) SAH7 protein [Arabidopsis thaliana]
- % Identity: 48.5
- Alignment Length: 163
- Location of Alignment in SEQ ID NO 300: from 1 to

35

146

Maximum Length Sequence:

related to:

Clone IDs:

31422

- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 301
  - Ceres seq\_id 1573861

45

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 302

50

206

- 5 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 197  
 - gi No. 2191138  
 - Description: (AF007269) A\_IG002N01.18 gene  
 10 product [Arabidopsis thaliana]  
 - % Identity: 28.1  
 - Alignment Length: 160  
 - Location of Alignment in SEQ ID NO 302: from 3 to 156
- 15 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 303  
 - Ceres seq\_id 1573863  
 - Location of start within SEQ ID NO 301: at 223 nt.
- 20 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 198  
 - gi No. 2191138  
 - Description: (AF007269) A\_IG002N01.18 gene  
 25 product [Arabidopsis thaliana]  
 - % Identity: 28.1  
 - Alignment Length: 160  
 - Location of Alignment in SEQ ID NO 303: from 1 to 104

- 35 (Ba) Polypeptide Activities: Arabidopsis specific gene,  
 dicot specific gene, plant specific gene.
- Maximum Length Sequence:  
 related to:  
 Clone IDs:  
 40916  
 40 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 304  
 - Ceres seq\_id 1574093

- 45 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 305  
 - Ceres seq\_id 1574094  
 - Location of start within SEQ ID NO 304: at 47 nt.

- 50 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)

207

- Zinc finger, C3HC4 type (RING finger)  
 - Location within SEQ ID NO 305: from 152 to 192
- aa.
- 5 (Dp) Related Amino Acid Sequences  
 - Alignment No. 199  
 - gi No. 3790593  
 - Description: (AF079185) RING-H2 finger protein  
 10 RHY1a [Arabidopsis thaliana]  
 - % Identity: 55.1  
 - Alignment Length: 49  
 - Location of Alignment in SEQ ID NO 305: from 148 to 196
- 15 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 306  
 - Ceres seq\_id 1574095  
 - Location of start within SEQ ID NO 304: at 56 nt.
- 20 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 - Zinc finger, C3HC4 type (RING finger)  
 - Location within SEQ ID NO 306: from 149 to 189
- aa.
- 25 (Dp) Related Amino Acid Sequences  
 - Alignment No. 200  
 - gi No. 3790593  
 - Description: (AF079185) RING-H2 finger protein  
 30 RHY1a [Arabidopsis thaliana]  
 - % Identity: 55.1  
 - Alignment Length: 49  
 - Location of Alignment in SEQ ID NO 306: from 145 to 193

Maximum Length Sequence:

related to:

Clone IDs:

35447

Public Genomic DNA:

gi No: 4263694

Predicted Exons:

INIT 79640 ... 79806 GENBANK

TERM 79888 ... 80164 GENBANK

INTR 79508 ... 79806 OCKHAMG-CDNA

INTR 79888 ... 80309 OCKHAMG-CDNA

INIT 79640 ... 79806 OCKHAMG-CDS

TERM 79888 ... 80164 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

208

- Pat. Appln. SEQ ID NO 307
- Ceres seq\_id 1580388
- Alternative transcription start site(s) located in SEQ

ID NO 307:

2, 3, 4, 5, 6, 7, 63, 94

5

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 308
  - Ceres seq\_id 1580389
  - Location of start within SEQ ID NO 307: at 133 nt.
  - Location of Signal Peptide Cleavage Site within SEQ
- ID NO 308: at 27 aa.

10

## (C) Nomination and Annotation of Domains within

- Predicted Polypeptide(s)
- Cystatin domain
- Location within SEQ ID NO 308: from 87 to 141 aa.

15

## (Dp) Related Amino Acid Sequences

- Alignment No. 201
  - gi No. 2204077
  - Description: (D85623) extracellular insoluble
- cystatin [Daucus carota]
- % Identity: 44.9
  - Alignment Length: 136
  - Location of Alignment in SEQ ID NO 308: from 6 to

20

25

141

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 309
  - Ceres seq\_id 1580390
  - Location of start within SEQ ID NO 307: at 142 nt.
  - Location of Signal Peptide Cleavage Site within SEQ
- ID NO 309: at 24 aa.

30

35

## (C) Nomination and Annotation of Domains within

- Predicted Polypeptide(s)
- Cystatin domain
- Location within SEQ ID NO 309: from 84 to 138 aa.

40

## (Dp) Related Amino Acid Sequences

- Alignment No. 202
  - gi No. 2204077
  - Description: (D85623) extracellular insoluble
- cystatin [Daucus carota]
- % Identity: 44.9
  - Alignment Length: 136
  - Location of Alignment in SEQ ID NO 309: from 3 to

45

138

50

Maximum Length Sequence:

209

related to:

Clone IDs:

22677

Public Genomic DNA:

gi No: 4972043

Predicted Exons:

SINGLE 18935 ... 19282 GENBANK

INTR

18856 ... 19527 OCKHAMG-CDNA

SINGLE

18935 ... 19282 OCKHAMG-CDS

gi No: 4972077

Predicted Exons:

INTR 60360 ... 61031 OCKHAMG-CDNA

SINGLE

60439 ... 60786 OCKHAMG-CDS

gi No: 5748495

Predicted Exons:

INTR 60360 ... 61031 OCKHAMG-CDNA

SINGLE

60439 ... 60786 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 310

- Ceres seq\_id 1582959

- Alternative transcription start site(s) located in SEQ
- ID NO 310:

2, 5, 6, 7, 58, 61

20

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 311
  - Ceres seq\_id 1582960
  - Location of start within SEQ ID NO 310: at 80 nt.
  - Location of Signal Peptide Cleavage Site within SEQ
- ID NO 311: at 19 aa.

30

35

## (C) Nomination and Annotation of Domains within

- Predicted Polypeptide(s)
- Dehydrins
- Location within SEQ ID NO 311: from 37 to 113 aa.

40

## (Dp) Related Amino Acid Sequences

- Alignment No. 203
  - gi No. 4972049
  - Description: (AL078470) glycine-rich protein like
- [Arabidopsis thaliana]
- % Identity: 100
  - Alignment Length: 115
  - Location of Alignment in SEQ ID NO 311: from 1 to

45

115

50

Maximum Length Sequence:

210

related to:  
Clone IDs:

15190  
34118

5 Public Genomic DNA:

gi No: 4582444

Predicted Exons:

INIT 39128 ... 39157 GENBANK  
INTR 39241 ... 39388 GENBANK  
INTR 39469 ... 39640 GENBANK  
INTR 39922 ... 40108 GENBANK  
TERM 40198 ... 40413 GENBANK

10

INTR 39945 ... 40108 OCKHAMG-CDNA  
INTR 40198 ... 40660 OCKHAMG-CDNA

15

INTR 39054 ... 39157 OCKHAMG-CDNA  
INTR 39241 ... 39388 OCKHAMG-CDNA  
INTR 39469 ... 39640 OCKHAMG-CDNA  
INTR 39922 ... 40108 OCKHAMG-CDNA  
INTR 40198 ... 40626 OCKHAMG-CDNA

20

INIT 39128 ... 39157 OCKHAMG-CDS  
INTR 39241 ... 39388 OCKHAMG-CDS  
INTR 39469 ... 39640 OCKHAMG-CDS  
INTR 39922 ... 40108 OCKHAMG-CDS  
TERM 40198 ... 40413 OCKHAMG-CDS

25

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 312

- Ceres seq\_id 1663221

- Alternative transcription start site(s) located in SEQ

ID NO 312:

-5,2,3,4,5,12,40,448

- Clone 15190 starts at 448 and ends at 1076 in cDNA.

- Clone 34118 starts at 1 and ends at 1042 in cDNA.

35

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 313

- Ceres seq\_id 1663222

- Location of start within SEQ ID NO 312: at 75 nt.

40

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Ribosomal protein S3, C-terminal domain.

- Location within SEQ ID NO 313: from 104 to 188

45

aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 204

- gi No. 133940

50

211

- Description: 40S RIBOSOMAL PROTEIN S3A (S1A)  
>gi|70851|pir|fR3XL3A ribosomal protein S3a - African clawed  
frog >gi|65091|emb|CAA40592| (X57322) ribosomal protein S1a  
[Xenopus laevis]

5 - % Identity: 80.1

- Alignment Length: 231

- Location of Alignment in SEQ ID NO 313: from 1 to

228

10 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 314

- Ceres seq\_id 1663223

- Location of start within SEQ ID NO 312: at 204 nt.

15

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Ribosomal protein S3, C-terminal domain.

- Location within SEQ ID NO 314: from 61 to 145 aa.

20

(Dp) Related Amino Acid Sequences

- Alignment No. 205

- gi No. 133940

- Description: 40S RIBOSOMAL PROTEIN S3A (S1A)

>gi|70851|pir|fR3XL3A ribosomal protein S3a - African clawed  
frog >gi|65091|emb|CAA40592| (X57322) ribosomal protein S1a  
[Xenopus laevis]

25

- % Identity: 80.1

- Alignment Length: 231

- Location of Alignment in SEQ ID NO 314: from 1 to

30 185

Maximum Length Sequence:

related to:

Clone IDs:

3996

2749

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 315

- Ceres seq\_id 1663275

- Alternative transcription start site(s) located in SEQ

ID NO 315:

3,4,5,6,7,8,9,15

- Clone 3996 starts at 1 and ends at 484 in cDNA.

- Clone 2749 starts at 7 and ends at 488 in cDNA.

45

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 316

- Ceres seq\_id 1663276

- Location of start within SEQ ID NO 315: at 75 nt.

50

212

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein S14p/S29e
- Location within SEQ ID NO 316: from 3 to 54 aa.

5

## (Dp) Related Amino Acid Sequences

- Alignment No. 206
- gi No. 4506717
- Description: ref|NP\_001023.1|PRPS29| ribosomal protein S29 >gi|266972|ep|P30054|RS29\_HUMAN 40S RIBOSOMAL PROTEIN S29 >gi|631884|pir||S30298|ribosomal protein S29 - rat >gi|362934|pir||S55919|ribosomal protein S29 - human protein S29 [Homo sapiens]
- % Identity: 72.2
- Alignment Length: 54
- Location of Alignment in SEQ ID NO 316: from 1 to

54

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 317
- Ceres seq\_id 1663277
- Location of start within SEQ ID NO 315: at 92 nt.

20

## (C) Nomination and Annotation of Domains within

## (Dp) Related Amino Acid Sequences

## Maximum Length Sequence:

- related to:
- Clone IDs: 114940 25068
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 318
- Ceres seq\_id 1665304
- Alternative transcription start site(s) located in SEQ ID NO 318:
  - 8,-3,-1,2,3,4,8,17,251
- Clone 114940 starts at 1 and ends at 743 in cDNA.
- Clone 25068 starts at 4 and ends at 760 in cDNA.

35

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 319
- Ceres seq\_id 1665305
- Location of start within SEQ ID NO 318: at 63 nt.

40

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 207
- gi No. 2129641

50

213

- Description: major latex protein type 1 - Arabidopsis thaliana >gi|1107493|emb|CAA63026| (X91960) major latex protein type1 [Arabidopsis thaliana]
- % Identity: 71.3
- Alignment Length: 150
- Location of Alignment in SEQ ID NO 319: from 1 to

5

150

## Maximum Length Sequence:

## related to:

## Clone IDs:

94673

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 320
- Ceres seq\_id 1709970
- Alternative transcription start site(s) located in SEQ ID NO 320:
  - 2,-1,11,41,42,43,44,45,46,47,48,51

15

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 321
- Ceres seq\_id 1709971
- Location of start within SEQ ID NO 320: at 276 nt.

20

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 208
- gi No. 2961300
- Description: (AJ225027) ribosomal protein L24 [Cicer arietinum]
- % Identity: 86.2
- Alignment Length: 160
- Location of Alignment in SEQ ID NO 321: from 1 to

30

107

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 322
- Ceres seq\_id 1709972
- Location of start within SEQ ID NO 320: at 424 nt.

35

40

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

## Maximum Length Sequence:

## related to:

## Clone IDs:

112835

(Ac) cDNA Polynucleotide Sequence

- Alignment No. 207
- Pat. Appln. SEQ ID NO 323

50

214

- Ceres seq\_id 1711273
- Alternative transcription start site(s) located in SEQ ID NO 323:

2,28,30,33,56,62,235

5

- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 324
  - Ceres seq\_id 1711274
  - Location of start within SEQ ID NO 323: at 92 nt.

10

- (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Helix-turn-helix
- Location within SEQ ID NO 324: from 87 to 141 aa.

15

- (Dp) Related Amino Acid Sequences

- Alignment No. 209
- gi No. 1632831
- Description: (Z49698) orf [Ricin communis]
- % Identity: 81
- Alignment Length: 142
- Location of Alignment in SEQ ID NO 324: from 1 to

142

25 Maximum Length Sequence:

related to:

Clone IDs:

17878

- (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 325
- Ceres seq\_id 1715423
- Alternative transcription start site(s) located in SEQ ID NO 325:

2,4,7,66

35

- (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 326
- Ceres seq\_id 1715424
- Location of start within SEQ ID NO 325: at 25 nt.

40

- (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 210
- gi No. 4337175
- Description: (AC006416) ESTs gb|T20589, gb|T04648, gb|AA597906, gb|T04111, gb|R84180, gb|R65428, gb|T44439, gb|T76570, gb|R90004, gb|T45020, gb|T42457, gb|T20921, gb|AA042762 and gb|AA720210 come from this gene. [Arabidopsis thaliana]
- % Identity: 44.8

50

215

- Alignment Length: 67
- Location of Alignment in SEQ ID NO 326: from 15 to 81

5

- (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 327
- Ceres seq\_id 1715425
- Location of start within SEQ ID NO 325: at 40 nt.

10

- (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 211
- gi No. 4337175
- Description: (AC006416) ESTs gb|T20589, gb|T04648, gb|AA597906, gb|T04111, gb|R84180, gb|R65428, gb|T44439, gb|T76570, gb|R90004, gb|T45020, gb|T42457, gb|T20921, gb|AA042762 and gb|AA720210 come from this gene. [Arabidopsis thaliana]
- % Identity: 44.8
- Alignment Length: 67
- Location of Alignment in SEQ ID NO 327: from 10 to 76

25

- (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 328
- Ceres seq\_id 1715426
- Location of start within SEQ ID NO 325: at 260 nt.

30

- (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 212
- gi No. 4337175
- Description: (AC006416) ESTs gb|T20589, gb|T04648, gb|AA597906, gb|T04111, gb|R84180, gb|R65428, gb|T44439, gb|T76570, gb|R90004, gb|T45020, gb|T42457, gb|T20921, gb|AA042762 and gb|AA720210 come from this gene. [Arabidopsis thaliana]
- % Identity: 56.1
- Alignment Length: 57
- Location of Alignment in SEQ ID NO 328: from 5 to

61

45

- (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot specific gene, plant specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

26744



216

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 329
- Ceres seq\_id 1715962
- Alternative transcription start site(s) located in SEQ ID NO 329:

2, 7, 9, 10, 26, 28, 30, 31, 32, 38, 39, 42, 43, 47, 48, 194; 247

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 330
- Ceres seq\_id 1715963
- Location of start within SEQ ID NO 329: at 87 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein L6
- Location within SEQ ID NO 330: from 12 to 194 aa.

## (Dp) Related Amino Acid Sequences

## - Alignment No. 213

- gi No. 1710581

- Description: 60S RIBOSOMAL PROTEIN L9

>gi|2129720|pir||S71255 ribosomal protein L9 - Arabidopsis thaliana >gi|1107489|emb|CAA63024| (X91958) 60S ribosomal protein L9 [Arabidopsis thaliana]

- % Identity: 93.3
- Alignment Length: 195
- Location of Alignment in SEQ ID NO 330: from 1 to

194

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 331
- Ceres seq\_id 1715964
- Location of start within SEQ ID NO 329: at 114 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein L6
- Location within SEQ ID NO 331: from 3 to 185 aa.

## (Dp) Related Amino Acid Sequences

## - Alignment No. 214

- gi No. 1710581

- Description: 60S RIBOSOMAL PROTEIN L9

>gi|2129720|pir||S71255 ribosomal protein L9 - Arabidopsis thaliana >gi|1107489|emb|CAA63024| (X91958) 60S ribosomal protein L9 [Arabidopsis thaliana]

- % Identity: 93.3

- Alignment Length: 195

- Location of Alignment in SEQ ID NO 331: from 1 to

185

217

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 332
- Ceres seq\_id 1715965
- Location of start within SEQ ID NO 329: at 369 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein L6
- Location within SEQ ID NO 332: from 1 to 100 aa.

## (Dp) Related Amino Acid Sequences

## - Alignment No. 215

- gi No. 1710581

- Description: 60S RIBOSOMAL PROTEIN L9

>gi|2129720|pir||S71255 ribosomal protein L9 - Arabidopsis thaliana >gi|1107489|emb|CAA63024| (X91958) 60S ribosomal protein L9 [Arabidopsis thaliana]

- % Identity: 93.3

- Alignment Length: 195

- Location of Alignment in SEQ ID NO 332: from 1 to

100

## Maximum Length Sequence:

related to:

Clone IDs:

41712

## Public Genomic DNA:

gi No: 4512656

## Predicted Exons:

INIT 106774... 106633 GENBANK

INTR 106277... 106194 GENBANK

INTR 105920... 105838 GENBANK

TERM 105749... 105630 GENBANK

35 INTR 106841... 106633 OCKHAMG-CDNA

INTR 106277... 106194 OCKHAMG-CDNA

INTR 105920... 105838 OCKHAMG-CDNA

INTR 105749... 105453 OCKHAMG-CDNA

40 INIT 106774... 106633 OCKHAMG-CDS

INTR 106277... 106194 OCKHAMG-CDS

INTR 105920... 105838 OCKHAMG-CDS

TERM 105749... 105630 OCKHAMG-CDS

gi No: 4559319

## Predicted Exons:

INIT 10855 ... 10714 GENBANK

INTR 10358 ... 10275 GENBANK

INTR 10001 ... 9919 GENBANK

TERM 9830 ... 9711 GENBANK

50 INTR 10922 ... 10714 OCKHAMG-CDNA

218

INTR 10358 ... 10275 OCKHAMG-CDNA  
 INTR 10001 ... 9919 OCKHAMG-CDNA  
 INTR 9830 ... 9534 OCKHAMG-CDNA

5

INIT 10855 ... 10714 OCKHAMG-CDS  
 INTR 10358 ... 10275 OCKHAMG-CDS  
 INTR 10001 ... 9919 OCKHAMG-CDS  
 TERM 9830 ... 9711 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

10

- Pat. Appln. SEQ ID NO 333  
 - Ceres seq\_id 1808584  
 - Alternative transcription start site(s) located in SEQ

ID NO 333:

34

15

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 334

- Ceres seq\_id 1808585

- Location of start within SEQ ID NO 333: at 1 nt.

20

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Helix-turn-helix

- Location within SEQ ID NO 334: from 110 to 164

25 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 216

- gi No. 1632831

- Description: (Z49698) orf [Ricinus communis]

- % Identity: 81

- Alignment Length: 142

- Location of Alignment in SEQ ID NO 334: from 24

to 165

35

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 335

- Ceres seq\_id 1808586

- Location of start within SEQ ID NO 333: at 70 nt.

40

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Helix-turn-helix

- Location within SEQ ID NO 335: from 87 to 141 aa.

45

(Dp) Related Amino Acid Sequences

- Alignment No. 217

- gi No. 1632831

- Description: (Z49698) orf [Ricinus communis]

- % Identity: 81

- Alignment Length: 142

50

219

- Location of Alignment in SEQ ID NO 335: from 1 to 142

Maximum Length Sequence:

related to:

Clone IDs:

42211

Public Genomic DNA:

gi No: 4914400

Predicted Exons:

INTR 90159 ... 89532 OCKHAMG-CDNA

SINGLE 90124 ... 89615 OCKHAMG-CDS

gi No: 4938473

Predicted Exons:

SINGLE 2368 ... 1859 GENBANK

INTR 2403 ... 1776 OCKHAMG-CDNA

SINGLE 2368 ... 1859 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 336

- Ceres seq\_id 1808591

- Alternative transcription start site(s) located in SEQ ID NO 336: 7,10

25

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 337

- Ceres seq\_id 1808592

- Location of start within SEQ ID NO 336: at 2 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 337: at 36 aa.

35

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 218

- gi No. 4926823

- Description: (AC004135) T17H7.8 [Arabidopsis

thaliana]

- % Identity: 46.1

- Alignment Length: 178

- Location of Alignment in SEQ ID NO 337: from 13

45 to 177

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 338

- Ceres seq\_id 1808593

- Location of start within SEQ ID NO 336: at 38 nt.

220

- Location of Signal Peptide Cleavage Site within SEQ ID NO 338: at 24 aa.

## (C) Nomination and Annotation of Domains within

## (Dp) Related Amino Acid Sequences

- Alignment No. 219
- gi No. 4926823
- Description: (AC004135) T17H7.8 [Arabidopsis

10 thaliana]

- % Identity: 46.1

- Alignment Length: 178

- Location of Alignment in SEQ ID NO 338: from 1 to

165

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 339
- Ceres seq\_id 1808594
- Location of start within SEQ ID NO 336: at 164 nt.

20

## (C) Nomination and Annotation of Domains within

## (Dp) Related Amino Acid Sequences

- Alignment No. 220
- gi No. 4926823
- Description: (AC004135) T17H7.8 [Arabidopsis

25 thaliana]

- % Identity: 46.1

- Alignment Length: 178

- Location of Alignment in SEQ ID NO 339: from 1 to

30 123

(Ba) Polypeptide Activities: Arabidopsis specific gene, dicot specific gene, plant specific gene.

## Maximum Length Sequence:

related to:

Clone IDs:

2217

40 (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 340

- Ceres seq\_id 1920563

- Alternative transcription start site(s) located in SEQ

ID NO 340:

6

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 341
- Ceres seq\_id 1920564

- Location of start within SEQ ID NO 340: at 1 nt.

50

221

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Oxidoreductase FAD/NAD-binding domain
- Location within SEQ ID NO 341: from 45 to 167 aa.

5

## (Dp) Related Amino Acid Sequences

- Alignment No. 221

- gi No. 729477

- Description: FERREDOXIN--NADP REDUCTASE PRECURSOR
- (FNR) >gi|320548|pir|IA44974 ferredoxin--NADP+ reductase (EC 1.18.1.2) precursor - common ice plant >gi|167256 (M25528) ferredoxin-NADP+ reductase precursor (fnrA; EC 1.6.7.1) [Mesembryanthemum

- % Identity: 89.8

- Alignment Length: 197

- Location of Alignment in SEQ ID NO 341: from 1 to

197

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 342

- Ceres seq\_id 1920565

- Location of start within SEQ ID NO 340: at 112 nt.

20

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Oxidoreductase FAD/NAD-binding domain
- Location within SEQ ID NO 342: from 8 to 130 aa.

25

## (Dp) Related Amino Acid Sequences

- Alignment No. 222

- gi No. 729477

- Description: FERREDOXIN--NADP REDUCTASE PRECURSOR
- (FNR) >gi|320548|pir|IA44974 ferredoxin--NADP+ reductase (EC 1.18.1.2) precursor - common ice plant >gi|167256 (M25528) ferredoxin-NADP+ reductase precursor (fnrA; EC 1.6.7.1) [Mesembryanthemum

- % Identity: 89.8

- Alignment Length: 197

- Location of Alignment in SEQ ID NO 342: from 1 to

40 160

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 343

- Ceres seq\_id 1920566

- Location of start within SEQ ID NO 340: at 118 nt.

45

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Oxidoreductase FAD/NAD-binding domain
- Location within SEQ ID NO 343: from 6 to 128 aa.

50

- 222
- (Dp) Related Amino Acid Sequences
- Alignment No. 223
  - gi No. 729477
- 5 (FNR) >gi|320548|pir||A44974 ferredoxin--NADP REDUCTASE PRECURSOR  
1.18.1.2) precursor - common ice plant >gi|167256 (M25528)  
ferredoxin-NADP+ reductase precursor (fnrA; EC 1.6.7.1)  
[Mesembryanthemum
- 10 - % Identity: 89.8  
- Alignment Length: 197  
- Location of Alignment in SEQ ID NO 343: from 1 to 158
- Maximum Length Sequence:  
related to:  
Clone IDs:  
13864
- Public Genomic DNA:  
gi No: 4309719
- 20 Predicted Exons:  
INIT 50987 ... 50773 GENBANK  
INTR 49622 ... 49352 GENBANK  
INTR 49257 ... 49162 GENBANK  
INTR 48896 ... 48786 GENBANK  
INTR 48686 ... 48377 GENBANK  
INTR 48133 ... 48043 GENBANK  
INTR 47924 ... 47591 GENBANK  
INTR 47475 ... 47327 GENBANK  
TERM 47084 ... 46550 GENBANK
- 30 (Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 344  
- Ceres seq\_id 1974419  
- Alternative transcription start site(s) located in SEQ ID NO 344:  
-5
- 35 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 345  
- Ceres seq\_id 1974420  
- Location of start within SEQ ID NO 344: at 81 nt.
- 40 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
- ATPases associated with various cellular activities (AAA)  
- Location within SEQ ID NO 345: from 454 to 664
- 45 aa.
- 50 (Dp) Related Amino Acid Sequences  
- Alignment No. 224  
- gi No. 4006905

- 223
- (Arabidopsis thaliana]
- Description: (Z99708) ATPase-like protein
  - % Identity: 84.3
  - Alignment Length: 633
  - Location of Alignment in SEQ ID NO 345: from 72 to 700
- 5 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 346  
- Ceres seq\_id 1974421  
- Location of start within SEQ ID NO 344: at 96 nt.
- 10 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
- ATPases associated with various cellular activities (AAA)  
- Location within SEQ ID NO 346: from 449 to 659
- 15 aa.
- 20 (Dp) Related Amino Acid Sequences  
- Alignment No. 225  
- gi No. 4006905
- 25 (Arabidopsis thaliana]
- Description: (Z99708) ATPase-like protein
  - % Identity: 84.3
  - Alignment Length: 633
  - Location of Alignment in SEQ ID NO 346: from 67 to 695
- 30 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 347  
- Ceres seq\_id 1974422  
- Location of start within SEQ ID NO 344: at 141 nt.
- 35 (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)  
- ATPases associated with various cellular activities (AAA)  
- Location within SEQ ID NO 347: from 434 to 644
- 40 aa.
- 45 (Dp) Related Amino Acid Sequences  
- Alignment No. 226  
- gi No. 4006905
- (Arabidopsis thaliana]
- Description: (Z99708) ATPase-like protein
  - % Identity: 84.3
  - Alignment Length: 633
  - Location of Alignment in SEQ ID NO 347: from 52 to 680
- 50 aa.

224

## Maximum Length Sequence:

related to:

Clone IDs:

36333

- 5 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 348  
 - Ceres seq\_id 1975983

- 10 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 349
- Ceres seq\_id 1975984
- Location of start within SEQ ID NO 348: at 90 nt.

- 15 (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 227
- gi No. 3688432
- Description: (AJ011705) 40S ribosomal protein S10
- [Lumbricus rubellus]
- % Identity: 55.3
- Alignment Length: 164
- Location of Alignment in SEQ ID NO 349: from 32

25 to 192

- (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 350
- Ceres seq\_id 1975985
- Location of start within SEQ ID NO 348: at 135 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 228
- gi No. 3688432
- Description: (AJ011705) 40S ribosomal protein S10
- [Lumbricus rubellus]
- % Identity: 55.3
- Alignment Length: 164
- Location of Alignment in SEQ ID NO 350: from 17

to 177

- (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 351
- Ceres seq\_id 1975986
- Location of start within SEQ ID NO 348: at 183 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 229
- gi No. 3688432
- Description: (AJ011705) 40S ribosomal protein S10
- [Lumbricus rubellus]
- % Identity: 55.3
- Alignment Length: 164
- Location of Alignment in SEQ ID NO 351: from 1 to

225

## Maximum Length Sequence:

related to:

Clone IDs:

24587

41072

- 5 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 352  
 - Ceres seq\_id 1976019  
 - Alternative transcription start site(s) located in SEQ ID NO 352: 13

(B) Polypeptide Sequence

(Dp) Related Amino Acid Sequences

- Alignment No. 230

- gi No. 3688432

- Description: (AJ011705) 40S ribosomal protein S10

[Lumbricus rubellus]

- % Identity: 55.3

- Alignment Length: 164

- Location of Alignment in SEQ ID NO 352: from 1 to 161

(Ba) Polypeptide Activities: Similar to glycine-rich RNA

binding ABA inducible protein activities.

Maximum Length Sequence:

related to:

Clone IDs:

24587

41072

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 352

- Ceres seq\_id 1976019

- Alternative transcription start site(s) located in SEQ ID NO 352: 13

- Clone 24587 starts at 13 and ends at in cDNA.

(B) Polypeptide Sequence

(Dp) Related Amino Acid Sequences

- Alignment No. 230

- gi No. 3688432

- Description: (AJ011705) 40S ribosomal protein S10

[Lumbricus rubellus]

- % Identity: 55.3

- Alignment Length: 164

- Location of Alignment in SEQ ID NO 353: from 1 to

40 161

(B) Polypeptide Sequence

(Dp) Related Amino Acid Sequences

- Alignment No. 230

- gi No. 3688432

- Description: (AJ011705) 40S ribosomal protein S10

[Lumbricus rubellus]

- % Identity: 55.3

- Alignment Length: 164

- Location of Alignment in SEQ ID NO 353: from 1 to

45 161

(B) Polypeptide Sequence

(Dp) Related Amino Acid Sequences

- Alignment No. 231

- gi No. 3688432

- Description: (AJ011705) 40S ribosomal protein S10

[Lumbricus rubellus]

- % Identity: 55.3

- Alignment Length: 164

- Location of Alignment in SEQ ID NO 353: from 1 to

- 226
- (Dp) Related Amino Acid Sequences
- Alignment No. 233
  - gi No. 1710750
  - Description: 30S RIBOSOMAL PROTEIN S21 (BS-B)
- 5 >gi11303814|dbj|BAA12470| (D84432) Yqex [Bacillus subtilis]
- 10 >gi11890063|dbj|BAA12082| (D83717) Yqex [Bacillus subtilis]
- >gi12634987|emb|CAB14483| (Z99117) ribosomal protein S21 [Bacillus subtilis]
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 355
  - Ceres seq\_id 1976022
  - Location of start within SEQ ID NO 352: at 287 nt.

## (C) Nomination and Annotation of Domains within

- Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 232
  - gi No. 3688432
  - Description: (AJ011705) 40S ribosomal protein S10 [Lumbricus rubellus]
  - % Identity: 55.3
  - Alignment Length: 164
  - Location of Alignment in SEQ ID NO 355: from 1 to
- 99

- 25 (Ba) Polypeptide Activities: Extensin like protein activities and glycine rich protein activities.

## Maximum Length Sequence:

related to:

Clone IDs:

117263

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 356
- Ceres seq\_id 1976673
- Alternative transcription start site(s) located in SEQ ID NO 356: -46,2,3,6,7

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 357
- Ceres seq\_id 1976674
- Location of start within SEQ ID NO 356: at 3 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 357: at 34 aa.

## (C) Nomination and Annotation of Domains within

- Predicted Polypeptide(s)
- Ribosomal protein S21
  - Location within SEQ ID NO 357: from 106 to 155

50 aa.

- 227
- (Dp) Related Amino Acid Sequences
- Alignment No. 233
  - gi No. 1710750
  - Description: 30S RIBOSOMAL PROTEIN S21 (BS-B)
- 5 >gi11303814|dbj|BAA12470| (D84432) Yqex [Bacillus subtilis]
- 10 >gi11890063|dbj|BAA12082| (D83717) Yqex [Bacillus subtilis]
- >gi12634987|emb|CAB14483| (Z99117) ribosomal protein S21 [Bacillus subtilis]
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 358
  - Ceres seq\_id 1976675
  - Location of start within SEQ ID NO 356: at 48 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 358: at 19 aa.
- 20
- (C) Nomination and Annotation of Domains within
- Predicted Polypeptide(s)
- Ribosomal protein S21
  - Location within SEQ ID NO 358: from 91 to 140 aa.
- 25
- (Dp) Related Amino Acid Sequences
- Alignment No. 234
  - gi No. 1710750
  - Description: 30S RIBOSOMAL PROTEIN S21 (BS-B)
- 30 >gi11303814|dbj|BAA12470| (D84432) Yqex [Bacillus subtilis]
- >gi11890063|dbj|BAA12082| (D83717) Yqex [Bacillus subtilis]
- >gi12634987|emb|CAB14483| (Z99117) ribosomal protein S21 [Bacillus subtilis]
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 359
  - Ceres seq\_id 1976676
  - Location of start within SEQ ID NO 356: at 192 nt.
- 35
- (C) Nomination and Annotation of Domains within
- Predicted Polypeptide(s)
- Ribosomal protein S21
  - Location within SEQ ID NO 359: from 43 to 92 aa.
- 50
- (Dp) Related Amino Acid Sequences
- Alignment No. 235
  - gi No. 1710750

228

- Description: 30S RIBOSOMAL PROTEIN S21 (BS-B)  
 >gi|1303814|dbj|BAA12470| (D84432) Yqex [Bacillus subtilis]  
 >gi|1890063|dbj|BAA12082| (D83717) Yqex [Bacillus subtilis]  
 >gi|2634987|emb|CAB14483| (Z99117) ribosomal protein S21  
 [Bacillus subtilis]  
 - % Identity: 38.5  
 - Alignment Length: 52  
 - Location of Alignment in SEQ ID NO 359: from 43

to 94

Maximum Length Sequence:  
 related to:  
 Clone IDs:  
 42333

Public Genomic DNA:

gi No: 6041810

Predicted Exons:

INTR 93582 ... 94052 OCKHAMG-CDNA  
 INTR 94169 ... 94771 OCKHAMG-CDNA

INIT 93630 ... 94052 OCKHAMG-CDS  
 TERM 94169 ... 94687 OCKHAMG-CDS

gi No: 6091711

Predicted Exons:

INTR 60479 ... 60949 OCKHAMG-CDNA  
 INTR 61066 ... 61668 OCKHAMG-CDNA

INIT 60527 ... 60949 OCKHAMG-CDS  
 TERM 61066 ... 61584 OCKHAMG-CDS

gi No: 6102641

Predicted Exons:

INTR 52165 ... 52635 OCKHAMG-CDNA  
 INTR 52752 ... 53354 OCKHAMG-CDNA

INIT 52213 ... 52635 OCKHAMG-CDS  
 TERM 52752 ... 53270 OCKHAMG-CDS

gi No: 6453849

Predicted Exons:

INTR 59432 ... 58964 OCKHAMG-CDNA  
 INTR 59847 ... 58243 OCKHAMG-CDNA

INIT 59386 ... 58964 OCKHAMG-CDS  
 TERM 59847 ... 58329 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 360  
 - Ceres seq\_id 2025186

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 361  
 - Ceres seq\_id 2025187

229

- Location of start within SEQ ID NO 360: at 1 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- D-isomer specific 2-hydroxyacid dehydrogenases  
 - Location within SEQ ID NO 361: from 93 to 295 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 236

- gi No. 5458864

- Description: (AJ248287) GLYCERATE DEHYDROGENASE  
 [Pyrococcus abyssi]

- % Identity: 42.8

- Alignment Length: 299

- Location of Alignment in SEQ ID NO 361: from 36  
 to 322

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 362

- Ceres seq\_id 2025188

- Location of start within SEQ ID NO 360: at 49 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- D-isomer specific 2-hydroxyacid dehydrogenases  
 - Location within SEQ ID NO 362: from 77 to 279 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 237

- gi No. 5458864

- Description: (AJ248287) GLYCERATE DEHYDROGENASE  
 [Pyrococcus abyssi]

- % Identity: 42.8

- Alignment Length: 299

- Location of Alignment in SEQ ID NO 362: from 20  
 to 306

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 363

- Ceres seq\_id 2025189

- Location of start within SEQ ID NO 360: at 70 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- D-isomer specific 2-hydroxyacid dehydrogenases  
 - Location within SEQ ID NO 363: from 70 to 272 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 238

- gi No. 5458864

230

- Description: (AJ248287) GLYCERATE DEHYDROGENASE  
[Pyrococcus abyssi]  
- % Identity: 42.8  
- Alignment Length: 299  
- Location of Alignment in SEQ ID NO 363: from 13  
to 299

Maximum Length Sequence:  
related to:

10

Clone IDs:  
267004

Public Genomic DNA:  
gi No: 2924730

Predicted Exons:  
INTR 8331 ... 8882

OCKHAMG-CDNA

SINGLE 8366 ... 8686

OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 364

- Ceres seq\_id 2025372

- Alternative transcription start site(s) located in SEQ

ID NO 364:

21

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 365

- Ceres seq\_id 2025373

- Location of start within SEQ ID NO 364: at 1 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 239

- gi No. 132825

- Description: 50S RIBOSOMAL PROTEIN CL25,

CHLOROPLAST PRECURSOR >gi|71308|pir||R5PM25 ribosomal protein

PsCL25 precursor, chloroplast - garden pea

>gi|20877|emb|CAA32187| (X14022) PsCL25 ribosomal preprotein

(AA -30 to 74) [Pisum sativum]

- % Identity: 56.1

- Alignment Length: 82

- Location of Alignment in SEQ ID NO 365: from 14

to 95

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 366

- Ceres seq\_id 2025374

- Location of start within SEQ ID NO 364: at 37 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 366: at 16 aa.

231

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 240

- gi No. 132825

- Description: 50S RIBOSOMAL PROTEIN CL25,

CHLOROPLAST PRECURSOR >gi|71308|pir||R5PM25 ribosomal protein

PsCL25 precursor, chloroplast - garden pea

>gi|20877|emb|CAA32187| (X14022) PsCL25 ribosomal preprotein

(AA -30 to 74) [Pisum sativum]

- % Identity: 56.1

- Alignment Length: 82

- Location of Alignment in SEQ ID NO 366: from 2 to

83

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 367

- Ceres seq\_id 2025375

- Location of start within SEQ ID NO 364: at 142 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 241

- gi No. 132825

- Description: 50S RIBOSOMAL PROTEIN CL25,

CHLOROPLAST PRECURSOR >gi|71308|pir||R5PM25 ribosomal protein

PsCL25 precursor, chloroplast - garden pea

>gi|20877|emb|CAA32187| (X14022) PsCL25 ribosomal preprotein

(AA -30 to 74) [Pisum sativum]

- % Identity: 56.1

- Alignment Length: 82

- Location of Alignment in SEQ ID NO 367: from 1 to

48

Maximum Length Sequence:

related to:

Clone IDs:

26930

Public Genomic DNA:

gi No: 5225383

Predicted Exons:

SINGLE 64539 ... 64276

GENBANK

INIT 64787 ... 64692

TERM 64620 ... 64276

gi No: 6136349

Predicted Exons:

INIT 126797... 126892

TERM 126964... 127308

OCKHAMG-CDS

OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence



232

- Pat. Appln. SEQ ID NO 368
- Ceres seq\_id 2025471
- Alternative transcription start site(s) located in SEQ ID NO 368:
- 95,96,97,98,100,101,102,105,106,110
- Clone 26930 starts at 97 and ends at 631 in cDNA.

5

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 369
- Ceres seq\_id 2025472
- Location of start within SEQ ID NO 368: at 1 nt.

10

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 242
- gi No. 5732082
- Description: (AF162444) contains similarity to Solanum lycopersicum (tomato) wound induced protein (GB:X59882) [Arabidopsis thaliana]
- % Identity: 100
- Alignment Length: 87
- Location of Alignment in SEQ ID NO 369: from 60 to 146

20

25

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 370
- Ceres seq\_id 2025473
- Location of start within SEQ ID NO 368: at 178 nt.

30

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 243
- gi No. 5732082
- Description: (AF162444) contains similarity to Solanum lycopersicum (tomato) wound induced protein (GB:X59882) [Arabidopsis thaliana]
- % Identity: 100
- Alignment Length: 87
- Location of Alignment in SEQ ID NO 370: from 1 to 87

40

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 371
- Ceres seq\_id 2025474
- Location of start within SEQ ID NO 368: at 214 nt.

45

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 245
- gi No. 5732082
- Description: (AF162444) contains similarity to Solanum lycopersicum (tomato) wound induced protein (GB:X59882) [Arabidopsis thaliana]
- % Identity: 100
- Alignment Length: 87

50

233

- Alignment No. 244
- gi No. 5732082
- Description: (AF162444) contains similarity to Solanum lycopersicum (tomato) wound induced protein (GB:X59882) [Arabidopsis thaliana]
- % Identity: 100
- Alignment Length: 87
- Location of Alignment in SEQ ID NO 371: from 1 to 75

5

10

## Maximum Length Sequence:

## related to:

## Clone IDs:

27681

## Public Genomic DNA:

- gi No: 6434227
- Predicted Exons:
- INTR 51354 ... 51417 OCKHAMG-CDNA
- INTR 51529 ... 51786 OCKHAMG-CDNA

15

20

25

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 372
- Ceres seq\_id 2025475
- Alternative transcription start site(s) located in SEQ ID NO 372:

30

1242,1251,1270,1271,1275,1276,1280,1287,1288,1290,1292,1293,1301,1302,1305,1307,1314,1318,1321,1325  
1326,1330

35

- Clone 27681 starts at 1242 and ends at 2029 in cDNA.

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 373
- Ceres seq\_id 2025476
- Location of start within SEQ ID NO 372: at 1 nt.

40

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 245
- gi No. 2739366
- Description: (AC002505) SF16 like protein [Arabidopsis thaliana]
- % Identity: 45
- Alignment Length: 373

45

50

234

to 418  
- Location of Alignment in SEQ ID NO 373: from 61

## (B) Polypeptide Sequence

- 5  
- Pat. Appln. SEQ ID NO 374  
- Ceres seq\_id 2025477  
- Location of start within SEQ ID NO 372: at 316 nt.

## (C) Nomination and Annotation of Domains within

## (Dp) Related Amino Acid Sequences

- 10  
- Alignment No. 246  
- gi No. 2739366  
- Description: (AC002505) SF16 like protein  
[Arabidopsis thaliana]

15  
- % Identity: 45

- Alignment Length: 373

- Location of Alignment in SEQ ID NO 374: from 1 to

313

## (B) Polypeptide Sequence

- 20  
- Pat. Appln. SEQ ID NO 375  
- Ceres seq\_id 2025478  
- Location of start within SEQ ID NO 372: at 364 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 375: at 21 aa.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 30  
- Alignment No. 247  
- gi No. 2739366  
- Description: (AC002505) SF16 like protein  
[Arabidopsis thaliana]

35  
- % Identity: 45

- Alignment Length: 373

- Location of Alignment in SEQ ID NO 375: from 1 to

297

## Maximum Length Sequence:

related to:

Clone IDs:

14564

Public Genomic DNA:

gi No: 4415905

Predicted Exons:

INTR	3235	...	3181	OCKHAMG-CDNA
INTR	2694	...	2619	OCKHAMG-CDNA
INTR	2509	...	2408	OCKHAMG-CDNA
INTR	2319	...	2171	OCKHAMG-CDNA
INTR	2076	...	1784	OCKHAMG-CDNA

50

235

INIT 3191 ... 3181 OCKHAMG-CDS

INTR 2694 ... 2619 OCKHAMG-CDS

INTR 2509 ... 2408 OCKHAMG-CDS

INTR 2319 ... 2171 OCKHAMG-CDS

TERM 2076 ... 1920 OCKHAMG-CDS

gi No: 4581138

Predicted Exons:

INTR 100911... 100857 OCKHAMG-CDNA

INTR 100370... 100295 OCKHAMG-CDNA

INTR 100185... 100084 OCKHAMG-CDNA

INTR 99995 ... 99847 OCKHAMG-CDNA

INTR 99752 ... 99460 OCKHAMG-CDNA

INIT 100867... 100857 OCKHAMG-CDS

INTR 100370... 100295 OCKHAMG-CDS

INTR 100185... 100084 OCKHAMG-CDS

INTR 99995 ... 99847 OCKHAMG-CDS

TERM 99752 ... 99596 OCKHAMG-CDS

20 (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 376

- Ceres seq\_id 2025524

- Alternative transcription start site(s) located in SEQ ID NO 376:  
-35,-26,-25,-2,3,4,5,6,7,8,9,15,16,17,32

25

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 377

- Ceres seq\_id 2025525

- Location of start within SEQ ID NO 376: at 47 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein L24e

- Location within SEQ ID NO 377: from 3 to 73 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 248

- gi No. 6094040

- Description: 60S RIBOSOMAL PROTEIN L24

>gi|2961300|emb|CAA12358| (AJ225027) ribosomal protein L24

[Cicer arietinum]

- % Identity: 88.3

- Alignment Length: 163

- Location of Alignment in SEQ ID NO 377: from 1 to

163

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 378

- Ceres seq\_id 2025526

- Location of start within SEQ ID NO 376: at 212 nt.

236

(C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 249  
 - gi No. 6094040  
 - Description: 60S RIBOSOMAL PROTEIN L24  
 >gi12961300|emb|CAA12358| (AJ225027) ribosomal protein L24  
 [Cicer arietinum]  
 - % Identity: 88.3  
 - Alignment Length: 163  
 - Location of Alignment in SEQ ID NO 378: from 1 to 108

15 Maximum Length Sequence:  
 related to:  
 Clone IDs:  
 20767  
 Public Genomic DNA:  
 gi No: 4519195  
 Predicted Exons:  
 INIT 50757 ... 50625 OCKHAMG-CDS  
 INTR 50001 ... 49847 OCKHAMG-CDS  
 INTR 49644 ... 49561 OCKHAMG-CDS  
 INTR 49391 ... 49236 OCKHAMG-CDS  
 INTR 49128 ... 49023 OCKHAMG-CDS  
 TERM 48810 ... 48662 OCKHAMG-CDS  
 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 379  
 - Ceres seq\_id 2025544  
 - Alternative transcription start site(s) located in SEQ ID NO 379:

35 656,672,680,685,701,703,706,709,713,717,728,736,741,743  
 - Clone 20767 starts at 656 and ends at 1623 in cDNA.

(B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 380  
 - Ceres seq\_id 2025545  
 - Location of start within SEQ ID NO 379: at 599 nt.  
 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 - Domain of unknown function  
 - Location within SEQ ID NO 380: from 134 to 247  
 aa.

50 (Dp) Related Amino Acid Sequences  
 - Alignment No. 250  
 - gi No. 1708292

SUBSTITUTE SHEET (RULE 26)

237

- Description: HEAT-RESPONSIVE PROTEIN 12  
 >gi1255116 (U50631) heat-responsive protein [Mus musculus]  
 - % Identity: 54.3  
 - Alignment Length: 129  
 - Location of Alignment in SEQ ID NO 380: from 119 to 247  
 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 381  
 - Ceres seq\_id 2025546  
 - Location of start within SEQ ID NO 379: at 1 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 381: at 50 aa.

15 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 - Zinc finger, C3HC4 type (RING finger)  
 - Location within SEQ ID NO 381: from 171 to 211  
 aa.

20 (Dp) Related Amino Acid Sequences  
 - Alignment No. 251  
 - gi No. 2435518  
 - Description: (AF024504) contains similarity to C3HC4-type zinc fingers [Arabidopsis thaliana]  
 - % Identity: 31.4  
 - Alignment Length: 221  
 - Location of Alignment in SEQ ID NO 381: from 35 to 211

30 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 382  
 - Ceres seq\_id 2025547  
 - Location of start within SEQ ID NO 379: at 16 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 382: at 45 aa.

35 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 - Zinc finger, C3HC4 type (RING finger)  
 - Location within SEQ ID NO 382: from 166 to 206  
 aa.

45 (Dp) Related Amino Acid Sequences  
 - Alignment No. 252  
 - gi No. 2435518  
 - Description: (AF024504) contains similarity to C3HC4-type zinc fingers [Arabidopsis thaliana]  
 - % Identity: 31.4  
 - Alignment Length: 221

SUBSTITUTE SHEET (RULE 26)

238

- Location of Alignment in SEQ ID NO 382: from 30 to 206

Maximum Length Sequence:

related to:

Clone IDs:

7802

Public Genomic DNA:

gi No: 4415928

Predicted Exons:

INTR 60672 ... 61210 OCKHAMG-CDNA

SINGLE 60712 ... 61119 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 383

- Ceres seq\_id 2026207

- Alternative transcription start site(s) located in SEQ

ID NO 383:

2,3,4,5

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 384

- Ceres seq\_id 2026208

- Location of start within SEQ ID NO 383: at 2 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 253

- gi No. 4757974

- Description: ref|NP\_004057.1|pCETN1| centrin, EF-

hand protein, 1 >gi|2493440|sp|Q12798|CAT2\_HUMAN CALTRACTIN,

ISOFORM 2 (CENTRIN) >gi|414993 (U03270) centrin [Homo

sapiens]

- % Identity: 34.3

- Alignment Length: 102

- Location of Alignment in SEQ ID NO 384: from 19

to 120

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 385

- Ceres seq\_id 2026209

- Location of start within SEQ ID NO 383: at 41 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 254

- gi No. 4757974

- Description: ref|NP\_004057.1|pCETN1| centrin, EF-

hand protein, 1 >gi|2493440|sp|Q12798|CAT2\_HUMAN CALTRACTIN,

239

ISOFORM 2 (CENTRIN) >gi|414993 (U03270) centrin [Homo sapiens]

- % Identity: 34.3

- Alignment Length: 102

- Location of Alignment in SEQ ID NO 385: from 6 to

5

107

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 386

- Ceres seq\_id 2026210

- Location of start within SEQ ID NO 383: at 62 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 255

- gi No. 4757974

- Description: ref|NP\_004057.1|pCETN1| centrin, EF-

hand protein, 1 >gi|2493440|sp|Q12798|CAT2\_HUMAN CALTRACTIN,

ISOFORM 2 (CENTRIN) >gi|414993 (U03270) centrin [Homo

sapiens]

- % Identity: 34.3

- Alignment Length: 102

- Location of Alignment in SEQ ID NO 386: from 1 to

25

100

Maximum Length Sequence:

Public Genomic DNA:

gi No: 4038029

Predicted Exons:

INIT 46188 ... 46125 GENBANK

TERM 46021 ... 45852 GENBANK

INTR 46258 ... 46125 OCKHAMG-CDNA

INTR 46021 ... 45663 OCKHAMG-CDNA

INIT 47157 ... 47119 OCKHAMG-CDS

INTR 46278 ... 46125 OCKHAMG-CDS

TERM 46021 ... 45852 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 387

- Ceres seq\_id 2026982

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 388

- Ceres seq\_id 2026983

- Location of start within SEQ ID NO 387: at 1 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 388: at 73 aa.

240

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- 5     - Gamma-thionins family  
      - Location within SEQ ID NO 388: from 74 to 120 aa.

(Dp) Related Amino Acid Sequences

- 10    - Alignment No. 256  
      - gi No. 1209258  
      - Description: (L31937) protease inhibitor II  
      [Brassica rapa]  
      - % Identity: 87  
      - Alignment Length: 77  
      - Location of Alignment in SEQ ID NO 388: from 44  
      to 120

(B) Polypeptide Sequence

- 15    - Pat. Appln. SEQ ID NO 389  
      - Ceres seq\_id 2026984  
      - Location of start within SEQ ID NO 387: at 34 nt.  
      - Location of Signal Peptide Cleavage Site within SEQ  
      ID NO 389: at 54 aa.

(C) Nomination and Annotation of Domains within

- 25    Predicted Polypeptide(s)  
      - Gamma-thionins family  
      - Location within SEQ ID NO 389: from 63 to 109 aa.

(Dp) Related Amino Acid Sequences

- 30    - Alignment No. 257  
      - gi No. 1209258  
      - Description: (L31937) protease inhibitor II  
      [Brassica rapa]  
      - % Identity: 87  
      - Alignment Length: 77  
      - Location of Alignment in SEQ ID NO 389: from 33  
      to 109

(B) Polypeptide Sequence

- 40    - Pat. Appln. SEQ ID NO 390  
      - Ceres seq\_id 2026985  
      - Location of start within SEQ ID NO 387: at 130 nt.  
      - Location of Signal Peptide Cleavage Site within SEQ  
      ID NO 390: at 22 aa.

(C) Nomination and Annotation of Domains within

- 45    Predicted Polypeptide(s)  
      - Gamma-thionins family  
      - Location within SEQ ID NO 390: from 31 to 77 aa.

(Dp) Related Amino Acid Sequences

50

241

- 5     - Alignment No. 258  
      - gi No. 1209258  
      - Description: (L31937) protease inhibitor II  
      [Brassica rapa]

- 77    - % Identity: 87  
      - Alignment Length: 77  
      - Location of Alignment in SEQ ID NO 390: from 1 to

(B) Polypeptide Sequence

- 10    - Pat. Appln. SEQ ID NO 391  
      - Ceres seq\_id 2026986  
      - Location of start within SEQ ID NO 387: at 30 nt.

(C) Nomination and Annotation of Domains within

- 15    Predicted Polypeptide(s)  
      (Dp) Related Amino Acid Sequences

Maximum Length Sequence:

related to:

Clone IDs:

251127

Public Genomic DNA:

gi No: 4220640

Predicted Exons:

- |      |           |       |             |
|------|-----------|-------|-------------|
| INIT | 67833 ... | 68135 | OCKHAMG-CDS |
| INTR | 68253 ... | 68366 | OCKHAMG-CDS |
| INTR | 68457 ... | 68552 | OCKHAMG-CDS |
| INTR | 68720 ... | 68794 | OCKHAMG-CDS |
| INTR | 68898 ... | 68965 | OCKHAMG-CDS |
| INTR | 69056 ... | 69066 | OCKHAMG-CDS |
| INTR | 69422 ... | 69624 | OCKHAMG-CDS |
| INTR | 69858 ... | 70202 | OCKHAMG-CDS |
| INTR | 70315 ... | 70428 | OCKHAMG-CDS |
| INTR | 70516 ... | 70611 | OCKHAMG-CDS |
| INTR | 70923 ... | 70997 | OCKHAMG-CDS |
| INTR | 71103 ... | 71173 | OCKHAMG-CDS |
| TERM | 71737 ... | 71749 | OCKHAMG-CDS |

(Ac) cDNA Polynucleotide Sequence

- 40    - Pat. Appln. SEQ ID NO 392  
      - Ceres seq\_id 2027300  
      - Alternative transcription start site(s) located in SEQ  
      ID NO 392:

38,53,55,63,82,84,85,103,160

(B) Polypeptide Sequence

- 45    - Pat. Appln. SEQ ID NO 393  
      - Ceres seq\_id 2027301  
      - Location of start within SEQ ID NO 392: at 177 nt.

242

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 259
- gi No. 5020103
- Description: (AF148219) fibrillin [Nostoc

5

PCC8009]

- % Identity: 30.3

- Alignment Length: 122

- Location of Alignment in SEQ ID NO 393: from 75

10

to 189

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO. 394

- Ceres seq\_id 2027302

- Location of start within SEQ ID NO 392: at 213 nt.

15

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 260

- gi No. 5020103

- Description: (AF148219) fibrillin [Nostoc

25

- % Identity: 30.3

- Alignment Length: 122

- Location of Alignment in SEQ ID NO 394: from 63

to 177

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 395

- Ceres seq\_id 2027303

- Location of start within SEQ ID NO 392: at 252 nt.

30

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 261

- gi No. 5020103

- Description: (AF148219) fibrillin [Nostoc

40

PCC8009]

- % Identity: 30.3

- Alignment Length: 122

- Location of Alignment in SEQ ID NO 395: from 50

45

to 164

## Maximum Length Sequence:

related to:

Clone IDs:

271260

Public Genomic DNA:

50

243

gi No: 6056186

## Predicted Exons:

INIT 66600 ... 66465 OCKHAMG-CDS  
TERM 66390 ... 66083 OCKHAMG-CDS

5

gi No: 6087919

## Predicted Exons:

INIT 31657 ... 31522 OCKHAMG-CDS  
TERM 31447 ... 31140 OCKHAMG-CDS

10

gi No: 6119514

## Predicted Exons:

INIT 32265 ... 32130 OCKHAMG-CDS  
TERM 32055 ... 31748 OCKHAMG-CDS

15

gi No: 6143857

## Predicted Exons:

INIT 34633 ... 34768 OCKHAMG-CDS  
TERM 34843 ... 35150 OCKHAMG-CDS

20

gi No: 6175136

## Predicted Exons:

INIT 34136 ... 34271 OCKHAMG-CDS  
TERM 34346 ... 34653 OCKHAMG-CDS

25

gi No: 6223634

## Predicted Exons:

INIT 34137 ... 34272 OCKHAMG-CDS  
TERM 34347 ... 34654 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 396

- Ceres seq\_id 2027375

30

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 397

- Ceres seq\_id 2027376

- Location of start within SEQ ID NO 396: at 2 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Bacterial mutT protein

- Location within SEQ ID NO 397: from 27 to 68 aa.

40

## (Dp) Related Amino Acid Sequences

- Alignment No. 262

- gi No. 2622420

- Description: (AE000895) mutator MutT protein

[Methanobacterium thermoautotrophicum]

- % Identity: 42.3

- Alignment Length: 98

- Location of Alignment in SEQ ID NO 397: from 13

to 109

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 398

50

244

- Ceres seq\_id 2027377
- Location of start within SEQ ID NO 396: at 17 nt.

## (C) Nomination and Annotation of Domains within

## 5 Predicted Polypeptide(s)

- Bacterial mutT protein
- Location within SEQ ID NO 398: from 22 to 63 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 263
- gi No. 2622420
- Description: (AE000895) mutator MutT protein [Methanobacterium thermoautotrophicum]
- % Identity: 42.3
- Alignment Length: 98
- Location of Alignment in SEQ ID NO 398: from 8 to 104

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 399
- Ceres seq\_id 2027378
- Location of start within SEQ ID NO 396: at 188 nt.

## (C) Nomination and Annotation of Domains within

## 25 Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 264
- gi No. 2622420
- Description: (AE000895) mutator MutT protein [Methanobacterium thermoautotrophicum]
- % Identity: 42.3
- Alignment Length: 98
- Location of Alignment in SEQ ID NO 399: from 1 to 47

## 35 Maximum Length Sequence:

## Public Genomic DNA:

gi No: 5868932

## Predicted Exons:

- |    |      |           |       |              |
|----|------|-----------|-------|--------------|
| 40 | INIT | 99857 ... | 96712 | GENBANK      |
|    | INTR | 96123 ... | 95883 | GENBANK      |
|    | TERM | 95642 ... | 95199 | GENBANK      |
| 45 | INTR | 96097 ... | 95883 | OCKHAMG-CDNA |
|    | INTR | 95642 ... | 95047 | OCKHAMG-CDNA |
|    | INIT | 96035 ... | 95883 | OCKHAMG-CDS  |
|    | TERM | 95642 ... | 95199 | OCKHAMG-CDS  |

## 50 (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 400

245

- Ceres seq\_id 2028729

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 401
- Ceres seq\_id 2028730
- Location of start within SEQ ID NO 400: at 1 nt.

## (C) Nomination and Annotation of Domains within

## 10 Predicted Polypeptide(s)

- C2 domain
- Location within SEQ ID NO 401: from 671 to 762 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 265
- gi No. 3142295
- Description: (AC002411) Strong similarity to phosphoribosylanthranilate transferase gb|D86180 from Pisum sativum. [Arabidopsis thaliana]
- % Identity: 45.6
- Alignment Length: 865
- Location of Alignment in SEQ ID NO 401: from 211 to 1050

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 402
- Ceres seq\_id 2028731
- Location of start within SEQ ID NO 400: at 199 nt.

## (C) Nomination and Annotation of Domains within

## 30 Predicted Polypeptide(s)

- C2 domain
- Location within SEQ ID NO 402: from 605 to 696 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 266
- gi No. 3142295
- Description: (AC002411) Strong similarity to phosphoribosylanthranilate transferase gb|D86180 from Pisum sativum. [Arabidopsis thaliana]
- % Identity: 45.6
- Alignment Length: 865
- Location of Alignment in SEQ ID NO 402: from 145 to 984

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 403
- Ceres seq\_id 2028732
- Location of start within SEQ ID NO 400: at 232 nt.

246

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)  
- C2 domain  
- Location within SEQ ID NO 403: from 594 to 685

5 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 267
- gi No. 3142295
- Description: (AC002411) Strong similarity to phosphoribosylanthranilate transferase gb|D86180 from *Pisum sativum*. [Arabidopsis thaliana]
- % Identity: 45.6
- Alignment Length: 865
- Location of Alignment in SEQ ID NO 403: from 134 to 973

20 Maximum Length Sequence:  
related to:

Clone IDs:  
9214

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 404
- Ceres seq\_id 2029079
- Alternative transcription start site(s) located in SEQ ID NO 404:  
-4,-1,2,3,4,8,9,10,11,14,15,16,17,18,20,23,24

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 405
- Ceres seq\_id 2029080
- Location of start within SEQ ID NO 404: at 3 nt.

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 268
- gi No. 132825
- Description: 50S RIBOSOMAL PROTEIN CL25, CHLOROPLAST PRECURSOR >gi|71308|pir||R5PM25 ribosomal protein PsCL25 precursor, chloroplast - garden pea
- >gi|20877|emb|CAA321871 (X14022) PsCL25 ribosomal preprotein (AA -30 to 74) [Pisum sativum]
- % Identity: 56.1
- Alignment Length: 82
- Location of Alignment in SEQ ID NO 405: from 14 to 95

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 406

247

- Ceres seq\_id 2029081
- Location of start within SEQ ID NO 404: at 39 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 406: at 16 aa.

5

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 269
- gi No. 132825
- Description: 50S RIBOSOMAL PROTEIN CL25, CHLOROPLAST PRECURSOR >gi|71308|pir||R5PM25 ribosomal protein PsCL25 precursor, chloroplast - garden pea
- >gi|20877|emb|CAA321871 (X14022) PsCL25 ribosomal preprotein (AA -30 to 74) [Pisum sativum]
- % Identity: 56.1
- Alignment Length: 82
- Location of Alignment in SEQ ID NO 406: from 2 to 83

20

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 407
- Ceres seq\_id 2029082
- Location of start within SEQ ID NO 404: at 144 nt.

25

## (C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 270
- gi No. 132825
- Description: 50S RIBOSOMAL PROTEIN CL25, CHLOROPLAST PRECURSOR >gi|71308|pir||R5PM25 ribosomal protein PsCL25 precursor, chloroplast - garden pea
- >gi|20877|emb|CAA321871 (X14022) PsCL25 ribosomal preprotein (AA -30 to 74) [Pisum sativum]
- % Identity: 56.1
- Alignment Length: 82
- Location of Alignment in SEQ ID NO 407: from 1 to 48

48

## Maximum Length Sequence:

## Public Genomic DNA:

gi No: 6434227

## Predicted Exons:

- |      |           |       |              |
|------|-----------|-------|--------------|
| INTR | 50076 ... | 50370 | OCKHAMG-CDNA |
| INTR | 50793 ... | 50874 | OCKHAMG-CDNA |
| INTR | 51163 ... | 51227 | OCKHAMG-CDNA |
| INTR | 51331 ... | 51417 | OCKHAMG-CDNA |
| INTR | 51529 ... | 51786 | OCKHAMG-CDNA |
| INIT | 50200 ... | 50370 | OCKHAMG-CDS  |

50



248

INTR 50793 ... 50874 OCKHAMG-CDS  
 INTR 51163 ... 51227 OCKHAMG-CDS  
 INTR 51331 ... 51417 OCKHAMG-CDS  
 TERM 51529 ... 51630 OCKHAMG-CDS

5 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 408  
 - Ceres seq\_id 2029225

10 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 409  
 - Ceres seq\_id 2029226  
 - Location of start within SEQ ID NO 408: at 125 nt.

15 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 271  
 - gi No. 1352051  
 - Description: ATP SYNTHASE D CHAIN, MITOCHONDRIAL  
 >gi|484646|pir|J50739 H+-transporting ATP synthase (EC  
 3.6.1.34) chain d precursor - rat >gi|286204|dbj|BAA02422|  
 (D13120) ATP synthase subunit d precursor [Rattus norvegicus]  
 - % Identity: 28.6  
 - Alignment Length: 145  
 - Location of Alignment in SEQ ID NO 409: from 15  
 to 154

20 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 410  
 - Ceres seq\_id 2029227  
 - Location of start within SEQ ID NO 408: at 194 nt.

25 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 272  
 - gi No. 1352051  
 - Description: ATP SYNTHASE D CHAIN, MITOCHONDRIAL  
 >gi|484646|pir|J50739 H+-transporting ATP synthase (EC  
 3.6.1.34) chain d precursor - rat >gi|286204|dbj|BAA02422|  
 (D13120) ATP synthase subunit d precursor [Rattus norvegicus]  
 - % Identity: 28.6  
 - Alignment Length: 145  
 - Location of Alignment in SEQ ID NO 410: from 1 to  
 131

30 Maximum Length Sequence:  
 Public Genomic DNA:  
 gi No: 6056185

35

40

45

50

249

Predicted Exons:  
 INTR 21555 ... 22692 OCKHAMG-CDNA  
 INTR 22891 ... 23339 OCKHAMG-CDNA

5 INTR 22502 ... 22692 OCKHAMG-CDNA  
 INTR 22891 ... 23338 OCKHAMG-CDNA

INIT 21706 ... 22692 OCKHAMG-CDS  
 TERM 22891 ... 23124 OCKHAMG-CDS

10 gi No: 6091703  
 Predicted Exons:  
 INTR 21555 ... 22692 OCKHAMG-CDNA  
 INTR 22891 ... 23339 OCKHAMG-CDNA

15 INTR 22502 ... 22692 OCKHAMG-CDNA  
 INTR 22891 ... 23338 OCKHAMG-CDNA

INIT 21706 ... 22692 OCKHAMG-CDS  
 TERM 22891 ... 23124 OCKHAMG-CDS

20 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 411  
 - Ceres seq\_id 2029422

25 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 412  
 - Ceres seq\_id 2029423  
 - Location of start within SEQ ID NO 411: at 152 nt.

30 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 - CONSTANS family zinc finger  
 - Location within SEQ ID NO 412: from 17 to 57 aa.

35 (Dp) Related Amino Acid Sequences  
 - Alignment No. 273  
 - gi No. 1076301  
 - Description: CONSTANS protein - Arabidopsis  
 thaliana >gi|1161514|emb|CAA64407| (X94937) CONSTANS protein  
 [Arabidopsis thaliana]  
 - % Identity: 49.2  
 - Alignment Length: 63  
 - Location of Alignment in SEQ ID NO 412: from 12  
 to 74

40

45

50 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 413  
 - Ceres seq\_id 2029424  
 - Location of start within SEQ ID NO 411: at 155 nt.

250

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
- CONSTANS family zinc finger  
- Location within SEQ ID NO 413: from 16 to 56 aa.

5

(Dp) Related Amino Acid Sequences  
- Alignment No. 274  
- gi No. 1076301  
- Description: CONSTANS protein - Arabidopsis  
thaliana >gi|1161514|emb|CAA64407| (X94937) CONSTANS protein  
[Arabidopsis thaliana]  
- % Identity: 49.2  
- Alignment Length: 63  
- Location of Alignment in SEQ ID NO 413: from 11  
to 73

15

(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 414  
- Ceres seq\_id 2029425  
- Location of start within SEQ ID NO 411: at 443 nt.

20

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 275  
- gi No. 1076301  
- Description: CONSTANS protein - Arabidopsis  
thaliana >gi|1161514|emb|CAA64407| (X94937) CONSTANS protein  
[Arabidopsis thaliana]  
- % Identity: 69.8  
- Alignment Length: 43  
- Location of Alignment in SEQ ID NO 414: from 260  
to 302

25

Maximum Length Sequence:  
Public Genomic DNA:  
gi No: 5832738  
Predicted Exons:  
INTR 56241 ... 56150 OCKHAMG-CDNA  
INTR 56041 ... 55864 OCKHAMG-CDNA  
INTR 55493 ... 55317 OCKHAMG-CDNA  
INTR 55216 ... 55085 OCKHAMG-CDNA  
INTR 54898 ... 54803 OCKHAMG-CDNA  
INTR 54725 ... 54609 OCKHAMG-CDNA  
INTR 54510 ... 54415 OCKHAMG-CDNA  
INTR 54288 ... 54176 OCKHAMG-CDNA  
INTR 54077 ... 53779 OCKHAMG-CDNA  
INIT 56181 ... 56150 OCKHAMG-CDS  
INTR 56041 ... 55864 OCKHAMG-CDS  
INTR 55493 ... 55317 OCKHAMG-CDS

30

35

40

45

50

251

INTR 55216 ... 55085 OCKHAMG-CDS  
INTR 54898 ... 54803 OCKHAMG-CDS  
INTR 54725 ... 54609 OCKHAMG-CDS  
INTR 54510 ... 54415 OCKHAMG-CDS  
INTR 54288 ... 54176 OCKHAMG-CDS  
TERM 54077 ... 53954 OCKHAMG-CDS

5

(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 415  
- Ceres seq\_id 2029806

10

(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 416  
- Ceres seq\_id 2029807  
- Location of start within SEQ ID NO 415: at 1 nt.

15

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
- Aminotransferase class IV  
- Location within SEQ ID NO 416: from 68 to 352 aa.

20

(Dp) Related Amino Acid Sequences  
- Alignment No. 276  
- gi No. 3540183  
- Description: (AC004122) Highly Similar to  
branched-chain amino acid aminotransferase [Arabidopsis  
thaliana]  
- % Identity: 53.8  
- Alignment Length: 318  
- Location of Alignment in SEQ ID NO 416: from 55  
to 371

25

30

(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 417  
- Ceres seq\_id 2029808  
- Location of start within SEQ ID NO 415: at 61 nt.

35

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
- Aminotransferase class IV  
- Location within SEQ ID NO 417: from 48 to 332 aa.

40

(Dp) Related Amino Acid Sequences  
- Alignment No. 277  
- gi No. 3540183  
- Description: (AC004122) Highly Similar to  
branched-chain amino acid aminotransferase [Arabidopsis  
thaliana]  
- % Identity: 53.8  
- Alignment Length: 318

45

50

- 252  
- Location of Alignment in SEQ ID NO 417: from 35 to 351
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 418  
- Ceres seq\_id 2029809  
- Location of start within SEQ ID NO 415: at 163 nt.
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
- Aminotransferase class IV  
- Location within SEQ ID NO 418: from 14 to 298 aa.
- (Dp) Related Amino Acid Sequences  
- Alignment No. 278  
- gi No. 3540183  
- Description: (AC004122) Highly similar to branched-chain amino acid aminotransferase [Arabidopsis thaliana]  
- % Identity: 53.8  
- Alignment Length: 318  
- Location of Alignment in SEQ ID NO 418: from 1 to 317
- Maximum Length Sequence:  
Public Genomic DNA:  
gi No: 5822667  
Predicted Exons:  
INTR 59013 ... 59999 OCKHAMG-CDNA  
SINGLE 59071 ... 59769 OCKHAMG-CDS  
gi No: 6041831  
Predicted Exons:  
INTR 53943 ... 54929 OCKHAMG-CDNA  
SINGLE 54001 ... 54699 OCKHAMG-CDS
- (Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 419  
- Ceres seq\_id 2030038
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 420  
- Ceres seq\_id 2030039  
- Location of start within SEQ ID NO 419: at 2 nt.
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 279

- 253  
- gi No. 4009495  
- Description: (AF068904) YlmG [Staphylococcus aureus]  
- % Identity: 31.7  
- Alignment Length: 63  
- Location of Alignment in SEQ ID NO 420: from 177 to 239
- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 421  
- Ceres seq\_id 2030040  
- Location of start within SEQ ID NO 419: at 59 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 421: at 19 aa.
- (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 280  
- gi No. 4009495  
- Description: (AF068904) YlmG [Staphylococcus aureus]  
- % Identity: 31.7  
- Alignment Length: 63  
- Location of Alignment in SEQ ID NO 421: from 158 to 220

- Maximum Length Sequence:  
Public Genomic DNA:  
gi No: 4539290  
Predicted Exons:  
INIT 6134 ... 6132 GENBANK  
INTR 6021 ... 5973 GENBANK  
INTR 5588 ... 5309 GENBANK  
INTR 5226 ... 5106 GENBANK  
TERM 4951 ... 4871 GENBANK
- INTR 6208 ... 6132 OCKHAMG-CDNA  
INTR 6021 ... 5973 OCKHAMG-CDNA  
INTR 5588 ... 5309 OCKHAMG-CDNA  
INTR 5226 ... 5106 OCKHAMG-CDNA  
INTR 4951 ... 4724 OCKHAMG-CDNA
- INIT 6134 ... 6132 OCKHAMG-CDS  
INTR 6021 ... 5973 OCKHAMG-CDS  
INTR 5588 ... 5309 OCKHAMG-CDS  
INTR 5226 ... 5106 OCKHAMG-CDS  
TERM 4951 ... 4871 OCKHAMG-CDS
- gi No: 4914454  
Predicted Exons:  
INTR 39557 ... 39481 OCKHAMG-CDNA

254  
 INTR 39370 ... 39322 OCKHAMG-CDNA  
 INTR 38937 ... 38658 OCKHAMG-CDNA  
 INTR 38575 ... 38455 OCKHAMG-CDNA  
 INTR 38300 ... 38073 OCKHAMG-CDNA

5

(Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 422  
 - Ceres seq\_id 2031778

10

(B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 423  
 - Ceres seq\_id 2031779  
 - Location of start within SEQ ID NO 422: at 3 nt.

15

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 281

- gi No. 3688432

- Description: (AJ011705) 40S ribosomal protein S10

[Lumbricus rubellus]

- % Identity: 53.1

- Alignment Length: 164

- Location of Alignment in SEQ ID NO 423: from 25

25 to 186

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 424

- Ceres seq\_id 2031780

- Location of start within SEQ ID NO 422: at 75 nt.

30

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 282

- gi No. 3688432

- Description: (AJ011705) 40S ribosomal protein S10

[Lumbricus rubellus]

- % Identity: 53.1

- Alignment Length: 164

- Location of Alignment in SEQ ID NO 424: from 1 to

162

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 425

- Ceres seq\_id 2031781

- Location of start within SEQ ID NO 422: at 210 nt.

50

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

255

(Dp) Related Amino Acid Sequences

- Alignment No. 283

- gi No. 3688432

- Description: (AJ011705) 40S ribosomal protein S10

[Lumbricus rubellus]

- % Identity: 53.1

- Alignment Length: 164

- Location of Alignment in SEQ ID NO 425: from 1 to

117

Maximum Length Sequence:

Public Genomic DNA:

gi No: 6434226

Predicted Exons:

INTR 17873 ... 17514 OCKHAMG-CDNA

INTR 17046 ... 16869 OCKHAMG-CDNA

INIT 17854 ... 17514 OCKHAMG-CDS

TERM 17046 ... 17028 OCKHAMG-CDS

20

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 426

- Ceres seq\_id 2032723

25

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 427

- Ceres seq\_id 2032724

- Location of start within SEQ ID NO 426: at 20 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 427: at 24 aa.

30

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Plant lipid transfer protein family

- Location within SEQ ID NO 427: from 28 to 115 aa.

35

(Dp) Related Amino Acid Sequences

- Alignment No. 284

- gi No. 899224

- Description: (X60318) E2 [Brassica napus]

>gi1384340|prf11905428A phospholipid transfer protein

[Brassica napus]

- % Identity: 79

- Alignment Length: 119

- Location of Alignment in SEQ ID NO 427: from 1 to

119

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 428

- Ceres seq\_id 2032725

50

256

- Location of start within SEQ ID NO 426: at 71 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Plant lipid transfer protein family
- Location within SEQ ID NO 428: from 11 to 98 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 285
- gi No. 899224
- Description: (X60318) E2 [Brassica napus]
- >gi|384340|prf|1905428A phospholipid transfer protein [Brassica napus]
- % Identity: 79
- Alignment Length: 119
- Location of Alignment in SEQ ID NO 428: from 1 to

102

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 429
- Ceres seq\_id 2032726
- Location of start within SEQ ID NO 426: at 1 nt.

20

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

## Maximum Length Sequence:

## Public Genomic DNA:

gi No: 6143856

## Predicted Exons:

INIT 30571 ... 30464 OCKHAMG-CDS  
TERM 30327 ... 30238 OCKHAMG-CDS

gi No: 6223633

## Predicted Exons:

INIT 30571 ... 30464 OCKHAMG-CDS  
TERM 30327 ... 30238 OCKHAMG-CDS

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 430
- Ceres seq\_id 2035536

40

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 431
- Ceres seq\_id 2035537
- Location of start within SEQ ID NO 430: at 1 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 431: at 20 aa.

50

257

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 286
- gi No. 3024192
- Description: NADH-UBIQUINONE OXIDOREDUCTASE MWFE (Y07708) NADH dehydrogenase [Mus musculus]
- % Identity: 31.5
- Alignment Length: 54
- Location of Alignment in SEQ ID NO 431: from 4 to

56

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 432
- Ceres seq\_id 2035538
- Location of start within SEQ ID NO 430: at 2 nt.

15

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 433
- Ceres seq\_id 2035539
- Location of start within SEQ ID NO 430: at 25 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 433: at 17 aa.

25

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 287
- gi No. 3024192
- Description: NADH-UBIQUINONE OXIDOREDUCTASE MWFE (Y07708) NADH dehydrogenase [Mus musculus]
- % Identity: 31.5
- Alignment Length: 54
- Location of Alignment in SEQ ID NO 433: from 1 to

48

## Maximum Length Sequence:

## Public Genomic DNA:

gi No: 6136357

## Predicted Exons:

INIT 51922 ... 51830 OCKHAMG-CDS  
INTR 51129 ... 51055 OCKHAMG-CDS  
INTR 50958 ... 50806 OCKHAMG-CDS  
INTR 50706 ... 50608 OCKHAMG-CDS  
TERM 50326 ... 50243 OCKHAMG-CDS

50

- 258
- gi No: 6143858
- Predicted Exons:
- INIT 51922 ... 51830 OCKHAMG-CDS
- INTR 51129 ... 51055 OCKHAMG-CDS
- INTR 50958 ... 50806 OCKHAMG-CDS
- INTR 50706 ... 50608 OCKHAMG-CDS
- TERM 50326 ... 50243 OCKHAMG-CDS
- gi No: 6449510
- Predicted Exons:
- INTR 9035 ... 9111 OCKHAMG-CDNA
- INTR 9208 ... 9360 OCKHAMG-CDNA
- INTR 9460 ... 9558 OCKHAMG-CDNA
- INTR 9840 ... 9992 OCKHAMG-CDNA
- 15 (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 434
- Ceres seq\_id 2035575
- 20 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 435
- Ceres seq\_id 2035576
- Location of start within SEQ ID NO 434: at 1 nt.
- 25 (C) Nomination and Annotation of Domains within
- (Dp) Related Amino Acid Sequences
- Alignment No. 288
- gi No. 6066609
- Description: (AJ009987) chloroplast channel
- forming outer membrane protein [Pisum sativum]
- % Identity: 54.2
- Alignment Length: 177
- Location of Alignment in SEQ ID NO 435: from 1 to
- 35 167
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 436
- Ceres seq\_id 2035577
- Location of start within SEQ ID NO 434: at 13 nt.
- 40 (C) Nomination and Annotation of Domains within
- Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 289
- gi No. 6066609
- Description: (AJ009987) chloroplast channel
- forming outer membrane protein [Pisum sativum]
- % Identity: 54.2
- Alignment Length: 177
- 50
- 259
- Location of Alignment in SEQ ID NO 436: from 1 to
- 163
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 437
- Ceres seq\_id 2035578
- Location of start within SEQ ID NO 434: at 40 nt.
- 10 (C) Nomination and Annotation of Domains within
- Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 290
- gi No. 6066609
- Description: (AJ009987) chloroplast channel
- forming outer membrane protein [Pisum sativum]
- % Identity: 54.2
- Alignment Length: 177
- Location of Alignment in SEQ ID NO 437: from 1 to
- 154
- Maximum Length Sequence:
- Public Genomic DNA:
- gi No: 6434213
- Predicted Exons:
- INIT 54056 ... 54085 OCKHAMG-CDS
- INTR 54178 ... 54325 OCKHAMG-CDS
- INTR 54401 ... 54572 OCKHAMG-CDS
- INTR 54807 ... 54993 OCKHAMG-CDS
- TERM 55079 ... 55291 OCKHAMG-CDS
- 30 (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 438
- Ceres seq\_id 2036457
- 35 (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 439
- Ceres seq\_id 2036458
- Location of start within SEQ ID NO 438: at 1 nt.
- 40 (C) Nomination and Annotation of Domains within
- Predicted Polypeptide(s)
- Ribosomal protein S3, C-terminal domain.
- Location within SEQ ID NO 439: from 104 to 188
- 45 aa.
- (Dp) Related Amino Acid Sequences
- Alignment No. 291
- gi No. 133940
- Description: 40S RIBOSOMAL PROTEIN S3A (S1A)
- >gi|70851|pir|IR3XL3A ribosomal protein S3a - African clawed

260

- 5 frog >gi|65091|emb|CAA40592| (X57322) ribosomal protein Sla  
[Xenopus laevis]  
- % Identity: 75.1  
- Alignment Length: 250  
- Location of Alignment in SEQ ID NO 439: from 1 to
- 10 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 440  
- Ceres seq\_id 2036459  
- Location of start within SEQ ID NO 438: at 130 nt.
- 15 (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
- Ribosomal protein S3, C-terminal domain.  
- Location within SEQ ID NO 440: from 61 to 145 aa.

## (Dp) Related Amino Acid Sequences

- 20 - Alignment No. 292  
- gi No. 133940  
- Description: 40S RIBOSOMAL PROTEIN S3A (S1A)  
>gi|70851|pir|IR3XL3A ribosomal protein S3a - African clawed  
frog >gi|65091|emb|CAA40592| (X57322) ribosomal protein Sla  
[Xenopus laevis]  
- % Identity: 75.1  
- Alignment Length: 250  
- Location of Alignment in SEQ ID NO 440: from 1 to

206

## Maximum Length Sequence:

- 30 Public Genomic DNA:  
gi No: 6434218  
Predicted Exons:  
INIT 85292 ... 85282 OCKHAMG-CDS  
INTR 84903 ... 84828 OCKHAMG-CDS  
INTR 84730 ... 84629 OCKHAMG-CDS  
INTR 84524 ... 84376 OCKHAMG-CDS  
INTR 84282 ... 84156 OCKHAMG-CDS  
TERM 84148 ... 84065 OCKHAMG-CDS

## (Ac) cDNA Polynucleotide Sequence

- 40 - Pat. Appln. SEQ ID NO 441  
- Ceres seq\_id 2036585

45

- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 442  
- Ceres seq\_id 2036586  
- Location of start within SEQ ID NO 441: at 1 nt.

50

261

- 5 (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
- Ribosomal protein L24e  
- Location within SEQ ID NO 442: from 3 to 73 aa.
- 10 (Dp) Related Amino Acid Sequences  
- Alignment No. 293  
- gi No. 6094040  
- Description: 60S RIBOSOMAL PROTEIN L24  
>gi|2961300|emb|CAA12358| (AJ225027) ribosomal protein L24  
[Cicer arietinum]  
- % Identity: 86.6  
- Alignment Length: 158  
- Location of Alignment in SEQ ID NO 442: from 1 to

15 156

## (B) Polypeptide Sequence

- 20 - Pat. Appln. SEQ ID NO 443  
- Ceres seq\_id 2036587  
- Location of start within SEQ ID NO 441: at 166 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- 25 - Alignment No. 294  
- gi No. 6094040  
- Description: 60S RIBOSOMAL PROTEIN L24  
>gi|2961300|emb|CAA12358| (AJ225027) ribosomal protein L24  
[Cicer arietinum]  
- % Identity: 86.6  
- Alignment Length: 158  
- Location of Alignment in SEQ ID NO 443: from 1 to

30 101

## Maximum Length Sequence:

- 35 Public Genomic DNA:  
gi No: 5668776  
Predicted Exons:  
INIT 19552 ... 19704 OCKHAMG-CDS  
TERM 19947 ... 20279 OCKHAMG-CDS

## (Ac) cDNA Polynucleotide Sequence

- 40 - Pat. Appln. SEQ ID NO 444  
- Ceres seq\_id 2039554

45

- (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 445  
- Ceres seq\_id 2039555  
- Location of start within SEQ ID NO 444: at 1 nt.

50

262

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 295
- gi No. 5882720
- Description: (AC008263) Similar to gb|D86180 phosphoribosylanthranilate transferase from Pisum sativum and contains 2 PF|00168 C2 (phospholipid binding) domains. ESTs gb|H76726, gb|T45544 and gb|N96377 come from this gene.
- [Arab...
- % Identity: 99.3
- Alignment Length: 151
- Location of Alignment in SEQ ID NO 445: from 1 to

151

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 446

- Ceres seq\_id 2039556

- Location of start within SEQ ID NO 444: at 136 nt.

20

(C) Nomination and Annotation of Domains within

(Dp) Related Amino Acid Sequences

- Alignment No. 296

- gi No. 5882720
- Description: (AC008263) Similar to gb|D86180 phosphoribosylanthranilate transferase from Pisum sativum and contains 2 PF|00168 C2 (phospholipid binding) domains. ESTs gb|H76726, gb|T45544 and gb|N96377 come from this gene.
- [Arab...
- % Identity: 99.3
- Alignment Length: 151
- Location of Alignment in SEQ ID NO 446: from 1 to

30

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 447

- Ceres seq\_id 2044285

- Location of start within SEQ ID NO 447: at 13 nt.

20

(C) Nomination and Annotation of Domains within

(Dp) Related Amino Acid Sequences

- Alignment No. 298

- gi No. 4512613
- Description: (AC004793) F28K20.12 [Arabidopsis thaliana]
- % Identity: 99.2
- Alignment Length: 122
- Location of Alignment in SEQ ID NO 449: from 1 to

30

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 447

- Ceres seq\_id 2044286

- Location of start within SEQ ID NO 447: at 16 nt.

40

(C) Nomination and Annotation of Domains within

(Dp) Related Amino Acid Sequences

- Alignment No. 299

- gi No. 4512613
- Description: (AC004793) F28K20.12 [Arabidopsis thaliana]
- % Identity: 99.2
- Alignment Length: 122
- Location of Alignment in SEQ ID NO 450: from 1 to

50

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 447

- Ceres seq\_id 2044287

- Location of start within SEQ ID NO 447: at 16 nt.

263

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 297
- gi No. 4512613
- Description: (AC004793) F28K20.12 [Arabidopsis thaliana]
- % Identity: 99.2
- Alignment Length: 122
- Location of Alignment in SEQ ID NO 448: from 1 to

5

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 297
- gi No. 4512613
- Description: (AC004793) F28K20.12 [Arabidopsis thaliana]
- % Identity: 99.2
- Alignment Length: 122
- Location of Alignment in SEQ ID NO 448: from 1 to

10

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 297
- gi No. 4512613
- Description: (AC004793) F28K20.12 [Arabidopsis thaliana]
- % Identity: 99.2
- Alignment Length: 122
- Location of Alignment in SEQ ID NO 448: from 1 to

15

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 298
- gi No. 4512613
- Description: (AC004793) F28K20.12 [Arabidopsis thaliana]
- % Identity: 99.2
- Alignment Length: 122
- Location of Alignment in SEQ ID NO 449: from 1 to

20

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 298
- gi No. 4512613
- Description: (AC004793) F28K20.12 [Arabidopsis thaliana]
- % Identity: 99.2
- Alignment Length: 122
- Location of Alignment in SEQ ID NO 449: from 1 to

25

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 298
- gi No. 4512613
- Description: (AC004793) F28K20.12 [Arabidopsis thaliana]
- % Identity: 99.2
- Alignment Length: 122
- Location of Alignment in SEQ ID NO 449: from 1 to

30

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 299
- gi No. 4512613
- Description: (AC004793) F28K20.12 [Arabidopsis thaliana]
- % Identity: 99.2
- Alignment Length: 122
- Location of Alignment in SEQ ID NO 450: from 1 to

35

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 299
- gi No. 4512613
- Description: (AC004793) F28K20.12 [Arabidopsis thaliana]
- % Identity: 99.2
- Alignment Length: 122
- Location of Alignment in SEQ ID NO 450: from 1 to

40

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 299
- gi No. 4512613
- Description: (AC004793) F28K20.12 [Arabidopsis thaliana]
- % Identity: 99.2
- Alignment Length: 122
- Location of Alignment in SEQ ID NO 450: from 1 to

45

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 299
- gi No. 4512613
- Description: (AC004793) F28K20.12 [Arabidopsis thaliana]
- % Identity: 99.2
- Alignment Length: 122
- Location of Alignment in SEQ ID NO 450: from 1 to

50



264

## Maximum Length Sequence:

related to:

Clone IDs:

5 157859

Public Genomic DNA:

gi No: 4519187

Predicted Exons:

INIT 64212 ... 64316 OCKHAMG-CDS  
 INTR 64435 ... 64599 OCKHAMG-CDS  
 INTR 64655 ... 64789 OCKHAMG-CDS  
 INTR 64869 ... 64952 OCKHAMG-CDS  
 INTR 65028 ... 65143 OCKHAMG-CDS  
 INTR 65245 ... 65346 OCKHAMG-CDS  
 INTR 65417 ... 65530 OCKHAMG-CDS  
 INTR 65621 ... 65677 OCKHAMG-CDS  
 INTR 65790 ... 65856 OCKHAMG-CDS  
 INTR 65929 ... 65982 OCKHAMG-CDS  
 INTR 66089 ... 66163 OCKHAMG-CDS  
 INTR 66358 ... 66593 OCKHAMG-CDS  
 INTR 66699 ... 66811 OCKHAMG-CDS  
 TERM 68086 ... 68098 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 451

- Ceres seq\_id 2048114

- Alternative transcription start site(s) located in SEQ

ID NO 451:

456

- Clone 157859 starts at 456 and ends at in cDNA.

30

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 452

- Ceres seq\_id 2048115

- Location of start within SEQ ID NO 451: at 1 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 452: at 34 aa.

35

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 300

- gi No. 4850384

- Description: (AC007357) F3F19.3 [Arabidopsis

thaliana]

- % Identity: 66.2

- Alignment Length: 201

- Location of Alignment in SEQ ID NO 452: from 95

to 295

45

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 453

50

265

- Ceres seq\_id 2048116

- Location of start within SEQ ID NO 451: at 19 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 453: at 28 aa.

5

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 301

- gi No. 4850384

- Description: (AC007357) F3F19.3 [Arabidopsis

thaliana]

- % Identity: 66.2

- Alignment Length: 201

- Location of Alignment in SEQ ID NO 453: from 89

to 289

15

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 454

- Ceres seq\_id 2048117

- Location of start within SEQ ID NO 451: at 22 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 454: at 27 aa.

20

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 302

- gi No. 4850384

- Description: (AC007357) F3F19.3 [Arabidopsis

thaliana]

- % Identity: 66.2

- Alignment Length: 201

- Location of Alignment in SEQ ID NO 454: from 88

to 288

35

Maximum Length Sequence:

Public Genomic DNA:

gi No: 4519194

Predicted Exons:

INTR 15272 ... 15482 OCKHAMG-CDNA

INTR 15570 ... 15615 OCKHAMG-CDNA

INTR 16443 ... 16834 OCKHAMG-CDNA

INIT 15350 ... 15482 OCKHAMG-CDS

INTR 15570 ... 15615 OCKHAMG-CDS

TERM 16443 ... 16671 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 455

- Ceres seq\_id 2048271

50

266

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 456
- Ceres seq\_id 2048272
- Location of start within SEQ ID NO 455: at 79 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein S16
- Location within SEQ ID NO 456: from 9 to 70 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 303
- gi No. 3861401
- Description: (AJ235273) 30S RIBOSOMAL PROTEIN S16
- Rickettsia prowazekii
- % Identity: 43.7
- Alignment Length: 103
- Location of Alignment in SEQ ID NO 456: from 1 to 102

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 457
- Ceres seq\_id 2048273
- Location of start within SEQ ID NO 455: at 139 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein S16
- Location within SEQ ID NO 457: from 1 to 50 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 304
- gi No. 3861401
- Description: (AJ235273) 30S RIBOSOMAL PROTEIN S16
- Rickettsia prowazekii
- % Identity: 43.7
- Alignment Length: 103
- Location of Alignment in SEQ ID NO 457: from 1 to 82

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 458
- Ceres seq\_id 2048274
- Location of start within SEQ ID NO 455: at 232 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Related Amino Acid Sequences
- Alignment No. 305

267

## (B) Polypeptide Sequence

- gi No. 3861401
- Description: (AJ235273) 30S RIBOSOMAL PROTEIN S16
- Rickettsia prowazekii
- % Identity: 43.7
- Alignment Length: 103
- Location of Alignment in SEQ ID NO 458: from 1 to 51

## Maximum Length Sequence:

## Public Genomic DNA:

gi No: 4519195

## Predicted Exons:

- INIT 48116 ... 48096 OCKHAMG-CDS
- INTR 48033 ... 47989 OCKHAMG-CDS
- INTR 47604 ... 47531 OCKHAMG-CDS
- INTR 47435 ... 47386 OCKHAMG-CDS
- TERM 47286 ... 47192 OCKHAMG-CDS

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 459
- Ceres seq\_id 2048331

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 460
- Ceres seq\_id 2048332
- Location of start within SEQ ID NO 459: at 1 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Domain of unknown function
- Location within SEQ ID NO 460: from 8 to 92 aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 306
- gi No. 5032215
- Description: ref|NP\_005827.1|pUK114|translational inhibitor protein p14.5
- gi|1717975|sp|P52758|UK14 HUMAN 14.5 KD TRANSLATIONAL INHIBITOR PROTEIN (P14.5) [UK114 ANTIGEN HOMOLOG]
- gi|1177435|emb|CAA64670| sapiens]
- % Identity: 58.8
- Alignment Length: 85
- Location of Alignment in SEQ ID NO 460: from 8 to 92

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 461
- Ceres seq\_id 2048333
- Location of start within SEQ ID NO 459: at 79 nt.

268

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Domain of unknown function
- Location within SEQ ID NO 461: from 1 to 66 aa.

5

## (Dp) Related Amino Acid Sequences

- Alignment No. 307
- gi No. 5032215
- Description: ref|NP\_005827.1|pUK114|
- translational inhibitor protein p14.5
- >gi|1717975|sp|P52758|UK14\_HUMAN 14.5 KD TRANSLATIONAL
- INHIBITOR PROTEIN (P14.5) (UK114 ANTIGEN HOMOLOG)
- >gi|1177435|emb|CAA64670| sapiens]
- % Identity: 58.8
- Alignment Length: 85
- Location of Alignment in SEQ ID NO 461: from 1 to

15

66

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 462
- Ceres seq\_id 2048334
- Location of start within SEQ ID NO 459: at 139 nt.

20

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- (Dp) Related Amino Acid Sequences
- Alignment No. 308
- gi No. 5032215
- Description: ref|NP\_005827.1|pUK114|
- translational inhibitor protein p14.5
- >gi|1717975|sp|P52758|UK14\_HUMAN 14.5 KD TRANSLATIONAL
- INHIBITOR PROTEIN (P14.5) (UK114 ANTIGEN HOMOLOG)
- >gi|1177435|emb|CAA64670| sapiens]
- % Identity: 58.8
- Alignment Length: 85
- Location of Alignment in SEQ ID NO 462: from 1 to

25

30

35

46

## Maximum Length Sequence:

## Public Genomic DNA:

gi No: 4589412

## Predicted Exons:

INIT	37105 ...	37091	OCKHAMG-CDS
INTR	36413 ...	36339	OCKHAMG-CDS
INTR	36224 ...	36176	OCKHAMG-CDS
INTR	35977 ...	35698	OCKHAMG-CDS
INTR	35603 ...	35471	OCKHAMG-CDS
TERM	35193 ...	35119	OCKHAMG-CDS

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 463

50

269

## - Ceres seq\_id 2048466

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 464
- Ceres seq\_id 2048467
- Location of start within SEQ ID NO 463: at 1 nt.

5

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- (Dp) Related Amino Acid Sequences
- Alignment No. 309
- gi No. 3688432
- Description: (AJ011705) 40S ribosomal protein S10
- [Lumbricus rubellus]
- % Identity: 50.6
- Alignment Length: 164
- Location of Alignment in SEQ ID NO 464: from 30

15

to 193

20

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 465
- Ceres seq\_id 2048468
- Location of start within SEQ ID NO 463: at 4 nt.

25

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- (Dp) Related Amino Acid Sequences
- Alignment No. 310
- gi No. 3688432
- Description: (AJ011705) 40S ribosomal protein S10
- [Lumbricus rubellus]
- % Identity: 50.6
- Alignment Length: 164
- Location of Alignment in SEQ ID NO 465: from 29

30

35

to 192

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 466
- Ceres seq\_id 2048469
- Location of start within SEQ ID NO 463: at 88 nt.

40

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- (Dp) Related Amino Acid Sequences
- Alignment No. 311
- gi No. 3688432
- Description: (AJ011705) 40S ribosomal protein S10
- [Lumbricus rubellus]
- % Identity: 50.6
- Alignment Length: 164

45

50

270  
- Location of Alignment in SEQ ID NO 466: from 1 to

164

Maximum Length Sequence:

5 Public Genomic DNA:

gi No: 4914400

Predicted Exons:

INTR 97410 ... 97833 OCKHAMG-CDNA

10 SINGLE 97525 ... 97797 OCKHAMG-CDS

gi No: 4938473

Predicted Exons:

SINGLE 9769 ... 10041 GENBANK

15 INTR 9654 ... 10078 OCKHAMG-CDNA

INTR 10721 ... 10880 OCKHAMG-CDNA

SINGLE 9769 ... 10041 OCKHAMG-CDS

20 (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 467

- Ceres seq\_id 2050485

25 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 468

- Ceres seq\_id 2050486

- Location of start within SEQ ID NO 467: at 2 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 312

- gi No. 2493089

- Description: ATP SYNTHASE F CHAIN, MITOCHONDRIAL

- % Identity: 29.7

- Alignment Length: 77

- Location of Alignment in SEQ ID NO 468: from 52

to 125

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 469

- Ceres seq\_id 2050487

- Location of start within SEQ ID NO 467: at 116 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 313

- gi No. 2493089

- Description: ATP SYNTHASE F CHAIN, MITOCHONDRIAL

271

- % Identity: 29.7

- Alignment Length: 77

- Location of Alignment in SEQ ID NO 469: from 14

to 87

Maximum Length Sequence:

Public Genomic DNA:

gi No: 4454022

Predicted Exons:

INIT 55210 ... 55024 GENBANK

TERM 54477 ... 54209 GENBANK

INTR 55270 ... 55024 OCKHAMG-CDNA

INTR 54477 ... 54299 OCKHAMG-CDNA

INIT 55210 ... 55024 OCKHAMG-CDS

TERM 54477 ... 54209 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 470

- Ceres seq\_id 2050708

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 471

- Ceres seq\_id 2050709

- Location of start within SEQ ID NO 470: at 61 nt.

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 314

- gi No. 2129641

- Description: major latex protein type 1 -

Arabidopsis thaliana &gt;gi|1107493|emb|CAA63026| (X91960) major

latex protein type1 [Arabidopsis thaliana]

- % Identity: 71.3

- Alignment Length: 150

- Location of Alignment in SEQ ID NO 471: from 1 to

40 150

Maximum Length Sequence:

Public Genomic DNA:

gi No: 4914454

Predicted Exons:

INTR 42446 ... 41456 OCKHAMG-CDS

INTR 40358 ... 40280 OCKHAMG-CDS

INTR 39551 ... 39481 OCKHAMG-CDS

INTR 39415 ... 39322 OCKHAMG-CDS

INTR 39280 ... 39268 OCKHAMG-CDS

INTR 38937 ... 38658 OCKHAMG-CDS

272

INTR 38575 ... 38455 OCKHAMG-CDS  
TERM 38300 ... 38220 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 472  
- Ceres seq\_id 2050901

5

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 473  
- Ceres seq\_id 2050902  
- Location of start within SEQ ID NO 472: at 3 nt.

10

(C) Nomination and Annotation of Domains within

(Dp) Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 315  
- gi No. 3688432  
- Description: (AJ011705) 40S ribosomal protein S10  
- % Identity: 53.1  
- Alignment Length: 147  
- Location of Alignment in SEQ ID NO 473: from 416  
to 560

20

25

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 474  
- Ceres seq\_id 2050903  
- Location of start within SEQ ID NO 472: at 123 nt.

30

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 316  
- gi No. 3688432  
- Description: (AJ011705) 40S ribosomal protein S10  
- % Identity: 53.1  
- Alignment Length: 147  
- Location of Alignment in SEQ ID NO 474: from 376  
to 520

35

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 475  
- Ceres seq\_id 2050904  
- Location of start within SEQ ID NO 472: at 678 nt.  
- Location of Signal Peptide Cleavage Site within SEQ  
ID NO 475: at 39 aa.

45

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences  
- Alignment No. 318  
- gi No. 4725952  
- Description: (AL049730) AIR1A-like protein  
[Arabidopsis thaliana]  
- % Identity: 100  
- Alignment Length: 117

50

273

(Dp) Related Amino Acid Sequences

- Alignment No. 317  
- gi No. 3688432  
- Description: (AJ011705) 40S ribosomal protein S10  
- % Identity: 53.1  
- Alignment Length: 147  
- Location of Alignment in SEQ ID NO 475: from 191  
to 335

5

to 335

Maximum Length Sequence:

Public Genomic DNA:

gi No: 4725940

Predicted Exons:

SINGLE 96089 ... 95736 GENBANK

15

INTR 96156 ... 95622 OCKHAMG-CDNA

SINGLE 96083 ... 95736 OCKHAMG-CDS

20

Predicted Exons:

SINGLE 96089 ... 95736 GENBANK

INTR 96156 ... 95622 OCKHAMG-CDNA

25

SINGLE 96083 ... 95736 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 476  
- Ceres seq\_id 2051325

30

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 477  
- Ceres seq\_id 2051326  
- Location of start within SEQ ID NO 476: at 68 nt.  
- Location of Signal Peptide Cleavage Site within SEQ  
ID NO 477: at 30 aa.

35

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

- Plant lipid transfer protein family  
- Location within SEQ ID NO 477: from 34 to 115 aa.

40

(Dp) Related Amino Acid Sequences

- Alignment No. 318  
- gi No. 4725952  
- Description: (AL049730) AIR1A-like protein

45

[Arabidopsis thaliana]  
- % Identity: 100  
- Alignment Length: 117

50

117 274  
- Location of Alignment in SEQ ID NO 477: from 1 to

5 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 478  
- Ceres seq\_id 2051327  
- Location of start within SEQ ID NO 476: at 74 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 478: at 28 aa.

10 (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
- Plant lipid transfer protein family  
- Location within SEQ ID NO 478: from 32 to 113 aa.

15 (Dp) Related Amino Acid Sequences  
- Alignment No. 319  
- gi No. 4725952  
- Description: (AL049730) AIR1A-like protein  
[Arabidopsis thaliana]  
- % Identity: 100  
- Alignment Length: 117  
- Location of Alignment in SEQ ID NO 478: from 1 to

20 115

25 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 479  
- Ceres seq\_id 2051328  
- Location of start within SEQ ID NO 476: at 288 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 479: at 61 aa.

30 (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences

35 Maximum Length Sequence:  
Public Genomic DNA:  
gi No: 4490717  
Predicted Exons:  
INIT 87492 ... 87378 GENBANK  
INTR 87034 ... 86928 GENBANK  
INTR 86683 ... 86641 GENBANK  
TERM 86445 ... 86402 GENBANK

40 45  
INTR 87555 ... 87378 OCKHAMG-CDNA  
INTR 87034 ... 86928 OCKHAMG-CDNA  
INTR 86683 ... 86641 OCKHAMG-CDNA  
INTR 86445 ... 86257 OCKHAMG-CDNA  
INIT 87492 ... 87378 OCKHAMG-CDS

50

275  
INTR 87034 ... 86928 OCKHAMG-CDS  
INTR 86683 ... 86641 OCKHAMG-CDS  
TERM 86445 ... 86402 OCKHAMG-CDS

5 (Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 480  
- Ceres seq\_id 2051633

10 (B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 481  
- Ceres seq\_id 2051634  
- Location of start within SEQ ID NO 480: at 1 nt.

15 (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 320  
- gi No. 1083282  
- Description: cytochrome-c oxidase (EC 1.9.3.1) -  
mouse >gi|567766|I06465| cytochrome c oxidase [Mus musculus]  
>gi|1094404|prf|2106151A cytochrome c oxidase:SUBUNIT=VIA  
[Mus musculus]  
- % Identity: 41.4  
- Alignment Length: 106  
- Location of Alignment in SEQ ID NO 481: from 22 to 120

20 25  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 482  
- Ceres seq\_id 2051635  
- Location of start within SEQ ID NO 480: at 64 nt.

30 (C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)  
(Dp) Related Amino Acid Sequences  
- Alignment No. 321  
- gi No. 1083282  
- Description: cytochrome-c oxidase (EC 1.9.3.1) -  
mouse >gi|567766|I06465| cytochrome c oxidase [Mus musculus]  
>gi|1094404|prf|2106151A cytochrome c oxidase:SUBUNIT=VIA  
[Mus musculus]  
- % Identity: 41.4  
- Alignment Length: 106  
- Location of Alignment in SEQ ID NO 482: from 1 to

35 40 45  
99  
Maximum Length Sequence:  
Public Genomic DNA:  
gi No: 4584531  
Predicted Exons:

276  
 INTR 56873 ... 56641 OCKHANG-CDNA  
 INTR 56627 ... 56451 OCKHANG-CDNA  
 SINGLE 56812 ... 56645 OCKHANG-CDS

5 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 483  
 - Ceres seq\_id 2051906

10 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 484  
 - Ceres seq\_id 2051907  
 - Location of start within SEQ ID NO 483: at 2 nt.

15 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 322  
 - gi No. 4850407  
 - Description: (AC007357) EST gb|T21221 comes from  
 this gene. [Arabidopsis thaliana]  
 - % Identity: 57.1  
 - Alignment Length: 35  
 - Location of Alignment in SEQ ID NO 484: from 37  
 to 70

20 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 485  
 - Ceres seq\_id 2051908  
 - Location of start within SEQ ID NO 483: at 62 nt.

25 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 323  
 - gi No. 4850407  
 - Description: (AC007357) EST gb|T21221 comes from  
 this gene. [Arabidopsis thaliana]  
 - % Identity: 57.1  
 - Alignment Length: 35  
 - Location of Alignment in SEQ ID NO 485: from 17  
 to 50

30 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 486  
 - Ceres seq\_id 2051909  
 - Location of start within SEQ ID NO 483: at 74 nt.

35 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 324  
 - gi No. 4850407  
 - Description: (AC007357) EST gb|T21221 comes from  
 this gene. [Arabidopsis thaliana]  
 - % Identity: 57.1  
 - Alignment Length: 35  
 - Location of Alignment in SEQ ID NO 486: from 13  
 to 46

40 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 487  
 - Ceres seq\_id 2052403  
 - Location of start within SEQ ID NO 487: at 56 nt.

45 (C) Nomination and Annotation of Domains within  
 Predicted Polypeptide(s)  
 (Dp) Related Amino Acid Sequences  
 - Alignment No. 325  
 - gi No. 1321731  
 - Description: (Z72439) major allergen Cor a 1  
 [Corylus avellana]  
 - % Identity: 36.3  
 - Alignment Length: 159  
 - Location of Alignment in SEQ ID NO 488: from 5 to 155

50 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 488  
 - Ceres seq\_id 2052404  
 - Location of start within SEQ ID NO 488: from 5 to 155 aa.

278

## Predicted Exons:

INTR 82618 ... 82845 OCKHAMG-CDNA  
 INTR 83226 ... 83301 OCKHAMG-CDNA  
 INTR 83389 ... 83561 OCKHAMG-CDNA  
 INTR 84124 ... 84385 OCKHAMG-CDNA  
 INIT 82730 ... 82845 OCKHAMG-CDS  
 INTR 83226 ... 83301 OCKHAMG-CDS  
 INTR 83389 ... 83561 OCKHAMG-CDS  
 TERM 84124 ... 84199 OCKHAMG-CDS

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 489  
 - Ceres seq\_id 2053545

15

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 490  
 - Ceres seq\_id 2053546  
 - Location of start within SEQ ID NO 489: at 113 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 490: at 36 aa.

20

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- MAPEG family  
 - Location within SEQ ID NO 490: from 59 to 136 aa.

25

## (Dp) Related Amino Acid Sequences

- Alignment No. 326  
 - gi No. 4758714  
 - Description: ref|NP\_004519.1|pMGST3| microsomal glutathione S-transferase 3 >gi|2583081 (AF026977) microsomal glutathione S-transferase 3 [Homo sapiens]  
 - % Identity: 41.9  
 - Alignment Length: 131  
 - Location of Alignment in SEQ ID NO 490: from 8 to 136

35

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 491  
 - Ceres seq\_id 2053547  
 - Location of start within SEQ ID NO 489: at 197 nt.

40

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- MAPEG family  
 - Location within SEQ ID NO 491: from 31 to 108 aa.

45

## (Dp) Related Amino Acid Sequences

- Alignment No. 327

50

279

- gi No. 4758714  
 - Description: ref|NP\_004519.1|pMGST3| microsomal glutathione S-transferase 3 >gi|2583081 (AF026977) microsomal glutathione S-transferase 3 [Homo sapiens]  
 - % Identity: 41.9  
 - Alignment Length: 131  
 - Location of Alignment in SEQ ID NO 491: from 1 to 108

108

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 492  
 - Ceres seq\_id 2053548  
 - Location of start within SEQ ID NO 489: at 329 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 492: at 22 aa.

15

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- MAPEG family  
 - Location within SEQ ID NO 492: from 1 to 64 aa.

20

## (Dp) Related Amino Acid Sequences

- Alignment No. 328  
 - gi No. 4758714  
 - Description: ref|NP\_004519.1|pMGST3| microsomal glutathione S-transferase 3 >gi|2583081 (AF026977) microsomal glutathione S-transferase 3 [Homo sapiens]  
 - % Identity: 41.9  
 - Alignment Length: 131  
 - Location of Alignment in SEQ ID NO 492: from 1 to 64

30

## Maximum Length Sequence:

related to:

Clone IDs:

2004

31675

266907

## Public Genomic DNA:

gi No: 4887737

## Predicted Exons:

INTR 7907 ... 8579 OCKHAMG-CDNA

INTR 7913 ... 8587 OCKHAMG-CDNA

INTR 7911 ... 8579 OCKHAMG-CDNA

INTR 7909 ... 8579 OCKHAMG-CDNA

INTR 7913 ... 8578 OCKHAMG-CDNA

INTR 7913 ... 8578 OCKHAMG-CDNA

INTR 7913 ... 8578 OCKHAMG-CDNA

INTR 7913 ... 8578 OCKHAMG-CDNA

INTR 7913 ... 8578 OCKHAMG-CDNA

INTR 7913 ... 8578 OCKHAMG-CDNA

INTR 7913 ... 8578 OCKHAMG-CDNA

INTR 7913 ... 8578 OCKHAMG-CDNA

INTR 7913 ... 8578 OCKHAMG-CDNA



280

INTR 7913 ... 8509 OCKHAMG-CDNA

SINGLE 7958 ... 8425 OCKHAMG-CDS

gi No: 5019264

Predicted Exons:

SINGLE 4143 ... 4610 GENBANK

5

INTR 4092 ... 4764 OCKHAMG-CDNA

INTR 4098 ... 4772 OCKHAMG-CDNA

INTR 4096 ... 4764 OCKHAMG-CDNA

INTR 4094 ... 4764 OCKHAMG-CDNA

INTR 4098 ... 4763 OCKHAMG-CDNA

INTR 4098 ... 4694 OCKHAMG-CDNA

20

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 493

- Ceres seq\_id 2053884

- Alternative transcription start site(s) located in SEQ

25 ID NO 493:

-47,-3,2,3,4,5,6,7,8,10,11,12,13,14,15,16,20,21,28,29

38,44,45,60,68

- Clone 2004 starts at 1 and ends at 677 in cDNA.

- Clone 31675 starts at 3 and ends at 607 in cDNA.

- Clone 266907 starts at 4 and ends at 677 in cDNA.

30

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 494

- Ceres seq\_id 2053885

- Location of start within SEQ ID NO 493: at 1 nt.

35

(C) Nomination and Annotation of Domains within

Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 329

- gi No. 2498177

- Description: BASP1 PROTEIN

- % Identity: 28

- Alignment Length: 164

- Location of Alignment in SEQ ID NO 494: from 2 to

45

159

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 495

- Ceres seq\_id 2053886

- Location of start within SEQ ID NO 493: at 52 nt.

50

281

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 330

- gi No. 2498177

- Description: BASP1 PROTEIN

- % Identity: 28

- Alignment Length: 164

- Location of Alignment in SEQ ID NO 495: from 1 to

142

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 496

- Ceres seq\_id 2053887

- Location of start within SEQ ID NO 493: at 172 nt.

15

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 331

- gi No. 2498177

- Description: BASP1 PROTEIN

- % Identity: 28

- Alignment Length: 164

- Location of Alignment in SEQ ID NO 496: from 1 to

102

## Maximum Length Sequence:

## Public Genomic DNA:

gi No: 4887738

## Predicted Exons:

INTR 45004 ... 44725 OCKHAMG-CDNA

INTR 44038 ... 43918 OCKHAMG-CDNA

INTR 43839 ... 43444 OCKHAMG-CDNA

35

INIT 44942 ... 44725 OCKHAMG-CDS

INTR 44038 ... 43918 OCKHAMG-CDS

TERM 43839 ... 43564 OCKHAMG-CDS

gi No: 5103850

## Predicted Exons:

INIT 44944 ... 44727 GENBANK

INTR 44040 ... 43920 GENBANK

TERM 43841 ... 43566 GENBANK

40

INTR 45006 ... 44727 OCKHAMG-CDNA

INTR 44040 ... 43920 OCKHAMG-CDNA

INTR 43841 ... 43446 OCKHAMG-CDNA

45

INIT 44944 ... 44727 OCKHAMG-CDS

INTR 44040 ... 43920 OCKHAMG-CDS

50

282

TERM 43841 ... 43566 OCKHAMG-CDS

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 497
- Ceres seq\_id 2053908

5

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 498
- Ceres seq\_id 2053909
- Location of start within SEQ ID NO 497: at 63 nt.

10

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 332
- gi No. 2245000
- Description: (Z97341) LET1 like protein
- % Identity: 31.4
- Alignment Length: 156
- Location of Alignment in SEQ ID NO 498: from 17 to 164

20

## [Arabidopsis thaliana]

25

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 499
- Ceres seq\_id 2053910
- Location of start within SEQ ID NO 497: at 168 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 499: at 23 aa.

30

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 333
- gi No. 2245000
- Description: (Z97341) LET1 like protein
- % Identity: 31.4
- Alignment Length: 156
- Location of Alignment in SEQ ID NO 499: from 1 to 129

40

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 500
- Ceres seq\_id 2053911
- Location of start within SEQ ID NO 497: at 216 nt.

45

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

50

283

- Alignment No. 334

- gi No. 2245000

[Arabidopsis thaliana]

- Description: (Z97341) LET1 like protein
- % Identity: 31.4
- Alignment Length: 156
- Location of Alignment in SEQ ID NO 500: from 1 to 113

5

## Maximum Length Sequence:

## Public Genomic DNA:

gi No: 4335711

## Predicted Exons:

- |      |           |       |              |
|------|-----------|-------|--------------|
| INTR | 46239 ... | 46470 | OCKHAMG-CDNA |
| INTR | 46624 ... | 46739 | OCKHAMG-CDNA |
| INTR | 47125 ... | 47428 | OCKHAMG-CDNA |
| INTR | 46285 ... | 46470 | OCKHAMG-CDS  |
| INTR | 46624 ... | 46739 | OCKHAMG-CDS  |
| TERM | 47125 ... | 47413 | OCKHAMG-CDS  |

20

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 501
- Ceres seq\_id 2056123

25

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 502
- Ceres seq\_id 2056124
- Location of start within SEQ ID NO 501: at 47 nt.

30

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Zinc finger, C3HC4 type (RING finger)
- Location within SEQ ID NO 502: from 150 to 190

35

aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 335
- gi No. 3790593
- Description: (AF079185) RING-H2 finger protein
- % Identity: 55.1
- Alignment Length: 49
- Location of Alignment in SEQ ID NO 502: from 146 to 194

40

## PHYA [Arabidopsis thaliana]

45

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 503
- Ceres seq\_id 2056125
- Location of start within SEQ ID NO 501: at 56 nt.

50

284

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Zinc finger, C3HC4 type (RING finger)
- Location within SEQ ID NO 503: from 147 to 187

5

aa.

## (Dp) Related Amino Acid Sequences

- Alignment No. 336
- gi No. 3790593
- Description: (AF079185) RING-H2 finger protein RHY1a [Arabidopsis thaliana]
- % Identity: 55.1
- Alignment Length: 49
- Location of Alignment in SEQ ID NO 503: from 143 to 191

15

## Maximum Length Sequence:

## Public Genomic DNA:

- gi No: 4432829
- Predicted Exons:
- INTR 51185 ... 51056 OCKHAMG-CDNA
- INTR 50590 ... 50253 OCKHAMG-CDNA

20

25

30

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 504
- Ceres seq\_id 2056245

35

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 505
- Ceres seq\_id 2056246
- Location of start within SEQ ID NO 504: at 1 nt.

40

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

## (Dp) Related Amino Acid Sequences

- Alignment No. 337
- gi No. 4938484
- Description: (AL078464) transcription factor-like protein [Arabidopsis thaliana]
- % Identity: 56.2
- Alignment Length: 661

50

285

- Location of Alignment in SEQ ID NO 505: from 4 to 644

644

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 506
- Ceres seq\_id 2056247
- Location of start within SEQ ID NO 504: at 55 nt.

5

## (C) Nomination and Annotation of Domains within

## (Dp) Related Amino Acid Sequences

- Alignment No. 338
- gi No. 4938484
- Description: (AL078464) transcription factor-like protein [Arabidopsis thaliana]
- % Identity: 56.2
- Alignment Length: 661
- Location of Alignment in SEQ ID NO 506: from 1 to 626

15

20

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 507
- Ceres seq\_id 2056248
- Location of start within SEQ ID NO 504: at 751 nt.

25

## (C) Nomination and Annotation of Domains within

## (Dp) Related Amino Acid Sequences

- Alignment No. 339
- gi No. 4938484
- Description: (AL078464) transcription factor-like protein [Arabidopsis thaliana]
- % Identity: 56.2
- Alignment Length: 661
- Location of Alignment in SEQ ID NO 507: from 1 to 394

30

35

## Maximum Length Sequence:

## related to:

## Clone IDs:

213322

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 508
- Ceres seq\_id 1941142

40

45

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 509
- Ceres seq\_id 1941143
- Location of start within SEQ ID NO 508: at 119 nt.

50

286

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- Ribosomal protein L24e
- Location within SEQ ID NO 509: from 3 to 73 aa.

5

(Dp) Related Amino Acid Sequences

- Alignment No. 340
  - gi No. 1710521
  - Description: 60S RIBOSOMAL PROTEIN L24
- >gi|1154859|emb|CAA63960| (X94296) L24 ribosomal protein  
[Hordeum vulgare]
- % Identity: 94.4
  - Alignment Length: 162
  - Location of Alignment in SEQ ID NO 509: from 1 to

15 161

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 510
- Ceres seq\_id 1941144
- Location of start within SEQ ID NO 508: at 284 nt.

20

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 341
  - gi No. 1710521
  - Description: 60S RIBOSOMAL PROTEIN L24
- >gi|1154859|emb|CAA63960| (X94296) L24 ribosomal protein  
[Hordeum vulgare]

30

- % Identity: 94.4
- Alignment Length: 162
- Location of Alignment in SEQ ID NO 510: from 1 to

106

35 Maximum Length Sequence:  
related to:

Clone IDs:

241379

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 511
- Ceres seq\_id 1942975
- Alternative transcription start site(s) located in SEQ ID NO 511:

29,30,31,32,33,36

45

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 512
- Ceres seq\_id 1942976
- Location of start within SEQ ID NO 511: at 1 nt.

50

287

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- Ribosomal protein S3, C-terminal domain.
- Location within SEQ ID NO 512: from 149 to 233

5 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 342
  - gi No. 133940
  - Description: 40S RIBOSOMAL PROTEIN S3A (S1A)
- >gi|70851|pir|IR3XL3A ribosomal protein S3a - African clawed  
frog >gi|65091|emb|CAA40592| (X57322) ribosomal protein S1a  
[Xenopus laevis]
- % Identity: 82.5
  - Alignment Length: 223
  - Location of Alignment in SEQ ID NO 512: from 46

to 268

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 513
- Ceres seq\_id 1942977
- Location of start within SEQ ID NO 511: at 136 nt.

20

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- Ribosomal protein S3, C-terminal domain.
- Location within SEQ ID NO 513: from 104 to 188

aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 343
  - gi No. 133940
  - Description: 40S RIBOSOMAL PROTEIN S3A (S1A)
- >gi|70851|pir|IR3XL3A ribosomal protein S3a - African clawed  
frog >gi|65091|emb|CAA40592| (X57322) ribosomal protein S1a  
[Xenopus laevis]
- % Identity: 82.5
  - Alignment Length: 223
  - Location of Alignment in SEQ ID NO 513: from 1 to

40 223

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 514
- Ceres seq\_id 1942978
- Location of start within SEQ ID NO 511: at 205 nt.

45

(C) Nomination and Annotation of Domains within  
Predicted Polypeptide(s)

- Ribosomal protein S3, C-terminal domain.
- Location within SEQ ID NO 514: from 81 to 165 aa.

50

- 288
- (Dp) Related Amino Acid Sequences
- Alignment No. 344
  - gi No. 133940
  - Description: 40S RIBOSOMAL PROTEIN S3A (S1A)
- 5 >gi|70851|pir||R3X13A ribosomal protein S3a - African clawed frog >gi|65091|emb|CAA40592| (X57322) ribosomal protein S1a [Xenopus laevis]
- % Identity: 82.5
  - Alignment Length: 223
  - Location of Alignment in SEQ ID NO 514: from 1 to 200
- Maximum Length Sequence:
- related to:
- Clone IDs:
- 289536
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 515
  - Ceres seq\_id 1944349
- 20
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 516
  - Ceres seq\_id 1944350
  - Location of start within SEQ ID NO 515: at 423 nt.
- 25
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- Ribosomal protein L24e
  - Location within SEQ ID NO 516: from 3 to 73 aa.
- 30
- (Dp) Related Amino Acid Sequences
- Alignment No. 345
  - gi No. 1710521
  - Description: 60S RIBOSOMAL PROTEIN L24
- 35 >gi|1154859|emb|CAA63960| (X94296) L24 ribosomal protein [Hordeum vulgare]
- % Identity: 94.4
  - Alignment Length: 162
  - Location of Alignment in SEQ ID NO 516: from 1 to 161
- 40
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 517
  - Ceres seq\_id 1944351
  - Location of start within SEQ ID NO 515: at 588 nt.
- 45
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 346
- 50

- 289
- gi No. 1710521
  - Description: 60S RIBOSOMAL PROTEIN L24
- >gi|1154859|emb|CAA63960| (X94296) L24 ribosomal protein [Hordeum vulgare]
- % Identity: 94.4
  - Alignment Length: 162
  - Location of Alignment in SEQ ID NO 517: from 1 to 106
- 10
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 518
  - Ceres seq\_id 1944352
  - Location of start within SEQ ID NO 515: at 736 nt.
- 15
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- (Dp) Related Amino Acid Sequences
- Alignment No. 347
  - gi No. 1362587
  - Description: spermatid-specific protein T2 precursor - longfin squid
  - % Identity: 55.8
  - Alignment Length: 52
  - Location of Alignment in SEQ ID NO 518: from 1 to 25
- 20
- Maximum Length Sequence:
- related to:
- Clone IDs:
- 291258
- (Ac) cDNA Polynucleotide Sequence
- Pat. Appln. SEQ ID NO 519
  - Ceres seq\_id 1964011
  - Alternative transcription start site(s) located in SEQ ID NO 519: 4,34,37,38,39,40,41
- 35
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 520
  - Ceres seq\_id 1964012
  - Location of start within SEQ ID NO 519: at 3 nt.
- 40
- (C) Nomination and Annotation of Domains within Predicted Polypeptide(s)
- Ribosomal protein S3, C-terminal domain.
  - Location within SEQ ID NO 520: from 150 to 234 aa.
- 45
- (Dp) Related Amino Acid Sequences
- Alignment No. 348
  - gi No. 1173253
- 50

290

- Description: 40S RIBOSOMAL PROTEIN S3

>gi1543317|pir||S41170 ribosomal protein S3 - mouse

>gi157728|emb|CAA35916| (X51536) ribosomal protein S3 (AA 1-243) [Rattus rattus] >gi1439522|emb|CAA54167| (X76772)

5 ribosomal protein S3 [Mus musculus]

- % Identity: 81.1

- Alignment Length: 227

- Location of Alignment in SEQ ID NO 520: from 47 to 273

10

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 521

- Ceres seq\_id 1964013

- Location of start within SEQ ID NO 519: at 141 nt.

15

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Ribosomal protein S3, C-terminal domain.

- Location within SEQ ID NO 521: from 104 to 188

20 aa.

(Dp) Related Amino Acid Sequences

- Alignment No. 349

- gi No. 1173253

- Description: 40S RIBOSOMAL PROTEIN S3

>gi1543317|pir||S41170 ribosomal protein S3 - mouse

>gi157728|emb|CAA35916| (X51536) ribosomal protein S3 (AA 1-243) [Rattus rattus] >gi1439522|emb|CAA54167| (X76772)

30 ribosomal protein S3 [Mus musculus]

- % Identity: 81.1

- Alignment Length: 227

- Location of Alignment in SEQ ID NO 521: from 1 to 227

35

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 522

- Ceres seq\_id 1964014

- Location of start within SEQ ID NO 519: at 210 nt.

40

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

- Ribosomal protein S3, C-terminal domain.

- Location within SEQ ID NO 522: from 81 to 165 aa.

45

(Dp) Related Amino Acid Sequences

- Alignment No. 350

- gi No. 1173253

- Description: 40S RIBOSOMAL PROTEIN S3

>gi1543317|pir||S41170 ribosomal protein S3 - mouse

>gi157728|emb|CAA35916| (X51536) ribosomal protein S3 (AA 1-

291

>gi1439522|emb|CAA54167| (X76772)

ribosomal protein S3 [Mus musculus]

- % Identity: 81.1

- Alignment Length: 227

- Location of Alignment in SEQ ID NO 522: from 1 to 204

5

Maximum Length Sequence:  
related to:

10 Clone IDs:

207986

224937

227089

229580

241662

275880

278084

15

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 523

- Ceres seq\_id 1983854

- Alternative transcription start site(s) located in SEQ ID NO 523:

- 1,2,6,30,31,32,33,34,36

- Clone 207986 starts at 31 and ends at 810 in cDNA.

- Clone 224937 starts at 31 and ends at 810 in cDNA.

- Clone 227089 starts at 33 and ends at 804 in cDNA.

- Clone 229580 starts at 31 and ends at 821 in cDNA.

- Clone 241662 starts at 34 and ends at 815 in cDNA.

- Clone 275880 starts at 1 and ends at 804 in cDNA.

- Clone 278084 starts at 31 and ends at 815 in cDNA.

30

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 524

- Ceres seq\_id 1983855

- Location of start within SEQ ID NO 523: at 315 nt.

35

(C) Nomination and Annotation of Domains within Predicted Polypeptide(s)

(Dp) Related Amino Acid Sequences

- Alignment No. 351

- gi No. 1710521

- Description: 60S RIBOSOMAL PROTEIN L24

>gi11154859|emb|CAA63960| (X94296) L24 ribosomal protein [Hordeum vulgare]

- % Identity: 94

- Alignment Length: 150

- Location of Alignment in SEQ ID NO 524: from 1 to 106

45

Maximum Length Sequence:  
related to:

50

292

## Clone IDs:

221977  
226126  
293001

- 5 (Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 525

- Ceres seq\_id 1990261  
- Alternative transcription start site(s) located in SEQ

## ID NO 525:

- 10 16,30,32  
- Clone 226126 starts at 30 and ends at 772 in cDNA.  
- Clone 293001 starts at 32 and ends at 810 in cDNA.

## (B) Polypeptide Sequence

- 15 - Pat. Appln. SEQ ID NO 526  
- Ceres seq\_id 2061972  
- Location of start within SEQ ID NO 525: at 144 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- Ribosomal protein L24e  
- Location within SEQ ID NO 526: from 3 to 73 aa.

## (Dp) Related Amino Acid Sequences

- 25 - Alignment No. 352  
- gi No. 1710521  
- Description: 60S RIBOSOMAL PROTEIN L24  
>gi1154859|emb|CAA63960| (X94296) L24 ribosomal protein  
[Hordeum vulgare]  
- % Identity: 94.4  
- Alignment Length: 162  
- Location of Alignment in SEQ ID NO 526: from 1 to  
161

## (B) Polypeptide Sequence

- 35 - Pat. Appln. SEQ ID NO 527  
- Ceres seq\_id 2061973  
- Location of start within SEQ ID NO 525: at 309 nt.

## (C) Nomination and Annotation of Domains within

## Predicted Polypeptide(s)

- (Dp) Related Amino Acid Sequences  
- Alignment No. 353

- 45 - gi No. 1710521  
- Description: 60S RIBOSOMAL PROTEIN L24  
>gi1154859|emb|CAA63960| (X94296) L24 ribosomal protein  
[Hordeum vulgare]  
- % Identity: 94.4  
- Alignment Length: 162  
- Location of Alignment in SEQ ID NO 527: from 1 to  
106

293

## Maximum Length Sequence:

## related to:

- 5 Clone IDs:

9458

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 528  
- Ceres seq\_id 1007802

- 10 - Alternative transcription start site(s) located in SEQ  
ID NO 528:  
2,3,4

## (B) Polypeptide Sequence

- 15 - Pat. Appln. SEQ ID NO 529  
- Ceres seq\_id 1007803  
- Location of start within SEQ ID NO 528: at 3 nt.

## (B) Polypeptide Sequence

- 20 - Pat. Appln. SEQ ID NO 530  
- Ceres seq\_id 1007804  
- Location of start within SEQ ID NO 528: at 2 nt.  
- Location of Signal Peptide Cleavage Site within SEQ  
ID NO 530: at 29 aa.

## (B) Polypeptide Sequence

- 25 - Pat. Appln. SEQ ID NO 531  
- Ceres seq\_id 1007805  
- Location of start within SEQ ID NO 528: at 29 nt.  
- Location of Signal Peptide Cleavage Site within SEQ  
ID NO 531: at 20 aa.

- 30 (Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot

- 35 specific gene.  
Maximum Length Sequence:

## related to:

## Clone IDs:

8458

## (Ac) cDNA Polynucleotide Sequence

- 40 - Pat. Appln. SEQ ID NO 532  
- Ceres seq\_id 1008556

- 45 - Alternative transcription start site(s) located in SEQ  
ID NO 532:  
80,103,134,135,136,137,138,139,152,153,157,166,168

## (B) Polypeptide Sequence

- 50 - Pat. Appln. SEQ ID NO 533  
- Ceres seq\_id 1008557

294

- Location of start within SEQ ID NO 532: at 173 nt.

(Ba) Polypeptide Activities: Arabidopsis specific gene,

dicot

5 specific gene.

Maximum Length Sequence:

related to:

10 Clone IDs:

8384

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 534

- Ceres seq\_id 1008628

15 - Alternative transcription start site(s) located in SEQ

ID NO 534:

3,28

(B) Polypeptide Sequence

20 - Pat. Appln. SEQ ID NO 535

- Ceres seq\_id 1008629

- Location of start within SEQ ID NO 534: at 41 nt.

(Ba) Polypeptide Activities: Arabidopsis specific

25 gene, dicot

specific gene, plant

Maximum Length Sequence:

30 related to:

Clone IDs:

7260

Public Genomic DNA:

gi No: 5733889

Predicted Exons:

INIT 21286 ... 21509 GENBANK

INTR 21606 ... 21832 GENBANK

TERM 21951 ... 21958 GENBANK

40 INIT 21286 ... 21509 OCKHAMG-CDS

INTR 21606 ... 21832 OCKHAMG-CDS

TERM 21951 ... 21958 OCKHAMG-CDS

gi No: 6041810

Predicted Exons:

INIT 34915 ... 35138 OCKHAMG-CDS

INTR 35235 ... 35461 OCKHAMG-CDS

TERM 35580 ... 35587 OCKHAMG-CDS

gi No: 6091711

Predicted Exons:

INIT 100255... 100478 OCKHAMG-CDS

INTR 100575... 100801 OCKHAMG-CDS

295

TERM 100920... 100927 OCKHAMG-CDS

gi No: 6102641

Predicted Exons:

INIT 91941 ... 92164 OCKHAMG-CDS

INTR 92261 ... 92487 OCKHAMG-CDS

TERM 92606 ... 92613 OCKHAMG-CDS

gi No: 6453849

Predicted Exons:

INIT 19658 ... 19435 OCKHAMG-CDS

INTR 19338 ... 19112 OCKHAMG-CDS

TERM 18993 ... 18986 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 536

- Ceres seq\_id 1009376

15 - Alternative transcription start site(s) located in SEQ

ID NO 536:

4,7,22,48,56,77

(B) Polypeptide Sequence

20 - Pat. Appln. SEQ ID NO 537

- Ceres seq\_id 1009377

- Location of start within SEQ ID NO 536: at 238 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 537: at 76 aa.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 538

- Ceres seq\_id 1009378

- Location of start within SEQ ID NO 536: at 334 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 538: at 44 aa.

(B) Polypeptide Sequence

35 - Pat. Appln. SEQ ID NO 539

- Ceres seq\_id 1009379

- Location of start within SEQ ID NO 536: at 245 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 539: at 47 aa.

40 (Ba) Polypeptide Activities: Arabidopsis specific gene,

dicot

specific gene, plant

specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

4489

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 540



296

- Ceres seq\_id 1011128
- Alternative transcription start site(s) located in SEQ ID NO 540: 6,7

5

- (B) Polypeptide Sequence
  - Pat. Appln. SEQ ID NO 541
  - Ceres seq\_id 1011129
  - Location of start within SEQ ID NO 540: at 41 nt.

10

- (B) Polypeptide Sequence
  - Pat. Appln. SEQ ID NO 542
  - Ceres seq\_id 1011130
  - Location of start within SEQ ID NO 540: at 59 nt.

15

- (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot

20

specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

42241

25

- (Ac) cDNA Polynucleotide Sequence
  - Pat. Appln. SEQ ID NO 543
  - Ceres seq\_id 1011718

30

- (B) Polypeptide Sequence
  - Pat. Appln. SEQ ID NO 544
  - Ceres seq\_id 1011719
  - Location of start within SEQ ID NO 543: at 3 nt.

35

- (B) Polypeptide Sequence
  - Pat. Appln. SEQ ID NO 545
  - Ceres seq\_id 1011720
  - Location of start within SEQ ID NO 543: at 78 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 545: at 21 aa.

40

45

- (B) Polypeptide Sequence
  - Pat. Appln. SEQ ID NO 546
  - Ceres seq\_id 1011721
  - Location of start within SEQ ID NO 543: at 102 nt.

50

- (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot
- specific gene, plant

297

Maximum Length Sequence:  
related to:

- 5 Clone IDs: 42200

- (Ac) cDNA Polynucleotide Sequence
  - Pat. Appln. SEQ ID NO 547
  - Ceres seq\_id 1011735

10

- (B) Polypeptide Sequence
  - Pat. Appln. SEQ ID NO 548
  - Ceres seq\_id 1011736
  - Location of start within SEQ ID NO 547: at 75 nt.

15

- (B) Polypeptide Sequence
  - Pat. Appln. SEQ ID NO 549
  - Ceres seq\_id 1011737
  - Location of start within SEQ ID NO 547: at 1 nt.
  - Location of Signal Peptide Cleavage Site within SEQ ID NO 549: at 37 aa.

20

- (B) Polypeptide Sequence
  - Pat. Appln. SEQ ID NO 550
  - Ceres seq\_id 1011738
  - Location of start within SEQ ID NO 547: at 2 nt.

25

- (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot
- specific gene, plant

30

specific gene.

Maximum Length Sequence:

related to:

- 35 Clone IDs: 42217

- (Ac) cDNA Polynucleotide Sequence
  - Pat. Appln. SEQ ID NO 551
  - Ceres seq\_id 1011755
  - Alternative transcription start site(s) located in SEQ ID NO 551: 34

40

45

- (B) Polypeptide Sequence
  - Pat. Appln. SEQ ID NO 552
  - Ceres seq\_id 1011756
  - Location of start within SEQ ID NO 551: at 1 nt.

50

- (B) Polypeptide Sequence
  - Pat. Appln. SEQ ID NO 553.

298

- Ceres seq\_id 1011757
- Location of start within SEQ ID NO 551: at 19 nt.

5 gene, dicot  
(Ba) Polypeptide Activities: Arabidopsis specific  
specific gene, plant

specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

42034

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 554
- Ceres seq\_id 1011832

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 555
- Ceres seq\_id 1011833
- Location of start within SEQ ID NO 554: at 97 nt.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 556
- Ceres seq\_id 1011834
- Location of start within SEQ ID NO 554: at 100 nt.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 557
- Ceres seq\_id 1011835
- Location of start within SEQ ID NO 554: at 106 nt.

30 (Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot

specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

41812

Public Genomic DNA:

gi No: 6434227

Predicted Exons:

45 INIT 65854 ... 65974 OCKHAMG-CDS  
TERM 66067 ... 66206 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 558
- Ceres seq\_id 1011907

(B) Polypeptide Sequence

299

- Pat. Appln. SEQ ID NO 559
- Ceres seq\_id 1011908
- Location of start within SEQ ID NO 558: at 218 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 559: at 30 aa.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 560
- Ceres seq\_id 1011909
- Location of start within SEQ ID NO 558: at 239 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 560: at 23 aa.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 561
- Ceres seq\_id 1011910
- Location of start within SEQ ID NO 558: at 3 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 561: at 17 aa.

20 (Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot

specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

41828

Public Genomic DNA:

gi No: 4159707

Predicted Exons:

35 INTR 15084 ... 15201 OCKHAMG-CDNA  
INTR 15295 ... 15324 OCKHAMG-CDNA  
INTR 15414 ... 15533 OCKHAMG-CDNA  
INTR 15648 ... 15956 OCKHAMG-CDNA

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 562
- Ceres seq\_id 1011911

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 563
- Ceres seq\_id 1011912
- Location of start within SEQ ID NO 562: at 194 nt.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 564
- Ceres seq\_id 1011913
- Location of start within SEQ ID NO 562: at 230 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 564: at 22 aa.

300

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 565  
 - Ceres seq\_id 1011914  
 - Location of start within SEQ ID NO 562: at 2 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 565: at 28 aa.

- (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot  
 specific gene, plant

- 15 Maximum Length Sequence:  
 related to:  
 Clone IDs:  
 41723  
 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 566  
 - Ceres seq\_id 1011954
- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 567  
 - Ceres seq\_id 1011955  
 - Location of start within SEQ ID NO 566: at 1 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 567: at 40 aa.

- 30 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 568  
 - Ceres seq\_id 1011956  
 - Location of start within SEQ ID NO 566: at 107 nt.

- 35 (Ba) Polypeptide Activities: Similar to DNAB Protein Activities

- Maximum Length Sequence:  
 related to:  
 Clone IDs:  
 41730

- Public Genomic DNA:  
 gi No: 4678371  
 Predicted Exons:  
 SINGLE 48996 ... 49184 GENBANK  
 SINGLE 48996 ... 49184 OCKHAMG-CDS  
 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 569  
 - Ceres seq\_id 1011960

301

- Alternative transcription start site(s) located in SEQ ID NO 569:  
 20,22,25,43

- 5 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 570  
 - Ceres seq\_id 1011961  
 - Location of start within SEQ ID NO 569: at 69 nt.

- 10 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 571  
 - Ceres seq\_id 1011962  
 - Location of start within SEQ ID NO 569: at 78 nt.

- 15 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 572  
 - Ceres seq\_id 1011963  
 - Location of start within SEQ ID NO 569: at 90 nt.

- 20 (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot  
 specific gene, plant specific gene, similar to serine rich protein activities.

- 25 Maximum Length Sequence:  
 related to:  
 Clone IDs:  
 38977

- 30 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 573  
 - Ceres seq\_id 1014075  
 - Alternative transcription start site(s) located in SEQ ID NO 573:  
 -3,-2,-1,2,3,4,5,6,7,8,10,15,103

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 574  
 - Ceres seq\_id 1014076  
 - Location of start within SEQ ID NO 573: at 1 nt.

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 575  
 - Ceres seq\_id 1014077  
 - Location of start within SEQ ID NO 573: at 84 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 48: at 22 aa.

- (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot

302

specific gene, plant

specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

37081

Public Genomic DNA:

gi No: 4159707

Predicted Exons:

INIT 15458 ... 15533

TERM 15648 ... 15751

OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 576

- Ceres seq\_id 1015865

- Alternative transcription start site(s) located in SEQ

ID NO 576:

-29

20

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 577

- Ceres seq\_id 1015866

- Location of start within SEQ ID NO 576: at 281 nt.

25

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 578

- Ceres seq\_id 1015867

- Location of start within SEQ ID NO 576: at 317 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 578: at 22 aa.

30

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 579

- Ceres seq\_id 1015868

- Location of start within SEQ ID NO 576: at 3 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 579: at 15 aa.

40

(Ba) Polypeptide Activities: Arabidopsis specific

gene, dicot

specific gene, plant

45

Maximum Length Sequence:

related to:

Clone IDs:

29375

Public Genomic DNA:

gi No: 4159707

Predicted Exons:

INTR 15158 ... 15201

OCKHAMG-CDNA

50

303

INTR 15449 ... 15533

INTR 15648 ... 15947

OCKHAMG-CDNA

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 580

- Ceres seq\_id 1021371

- Alternative transcription start site(s) located in SEQ

ID NO 580:

-74

10

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 581

- Ceres seq\_id 1021372

- Location of start within SEQ ID NO 580: at 3 nt.

15

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 582

- Ceres seq\_id 1021373

- Location of start within SEQ ID NO 580: at 54 nt.

20

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 583

- Ceres seq\_id 1021374

- Location of start within SEQ ID NO 580: at 90 nt.

- Location of Signal Peptide Cleavage Site within SEQ

ID NO 583: at 22 aa.

25

(Ba) Polypeptide Activities:

Arabidopsis specific gene,

dicot

specific gene, plant

30

specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

27064

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 584

- Ceres seq\_id 1022578

40

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 585

- Ceres seq\_id 1022579

- Location of start within SEQ ID NO 584: at 1 nt.

45

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 586

- Ceres seq\_id 1022580

- Location of start within SEQ ID NO 584: at 157 nt.

50

(B) Polypeptide Sequence

304

- Pat. Appln. SEQ ID NO 587
- Ceres seq\_id 1022581
- Location of start within SEQ ID NO 584: at 10 nt.

5 (Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot  
specific gene, plant

10 Maximum Length Sequence:

related to:  
Clone IDs:

23773

Public Genomic DNA:

gi No: 6136357

Predicted Exons:

INTR 51972 ... 51830 OCKHAMG-CDNA  
INTR 51129 ... 51055 OCKHAMG-CDNA  
INTR 50948 ... 50806 OCKHAMG-CDNA  
INTR 50706 ... 50608 OCKHAMG-CDNA  
INTR 50326 ... 50147 OCKHAMG-CDNA

gi No: 6143858

Predicted Exons:

INTR 51972 ... 51830 OCKHAMG-CDNA  
INTR 51129 ... 51055 OCKHAMG-CDNA  
INTR 50948 ... 50806 OCKHAMG-CDNA  
INTR 50706 ... 50608 OCKHAMG-CDNA  
INTR 50326 ... 50147 OCKHAMG-CDNA

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 588
- Ceres seq\_id 1024240

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 589
- Ceres seq\_id 1024241
- Location of start within SEQ ID NO 588: at 54 nt.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 590
- Ceres seq\_id 1024242
- Location of start within SEQ ID NO 588: at 66 nt.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 591
- Ceres seq\_id 1024243
- Location of start within SEQ ID NO 588: at 404 nt.

(Ba) Polypeptide Activities: Arabidopsis specific  
gene, dicot  
specific gene, plant

305

Maximum Length Sequence:  
related to:

Clone IDs:

2030

Public Genomic DNA:

gi No: 5441914

Predicted Exons:

SINGLE 4802 ... 4449 OCKHAMG-CDS

gi No: 5881519

Predicted Exons:

SINGLE 4802 ... 4449 GENBANK

SINGLE 4802 ... 4449 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 592

- Ceres seq\_id 1026562

- Alternative transcription start site(s) located in SEQ  
ID NO 592:

36,41,43,44,88,90,94,102

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 593

- Ceres seq\_id 1026563

- Location of start within SEQ ID NO 592: at 1 nt.

- Location of Signal Peptide Cleavage Site within SEQ  
ID NO 593: at 51 aa.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 594

- Ceres seq\_id 1026564

- Location of start within SEQ ID NO 592: at 109 nt.

- Location of Signal Peptide Cleavage Site within SEQ  
ID NO 594: at 15 aa.

(Ba) Polypeptide Activities: Similar to cytochrome c  
oxidase II activities.

Maximum Length Sequence:

related to:

Clone IDs:

20097

Public Genomic DNA:

gi No: 573893

Predicted Exons:

INTR 66948 ... 67429 OCKHAMG-CDNA

SINGLE 66986 ... 67324 OCKHAMG-CDS

gi No: 5748882

Predicted Exons:

- 306  
INTR 41416 ... 40935 OCKHAMG-CDNA
- 5  
SINGLE 41380 ... 41042 OCKHAMG-CDS  
gi No: 5801671  
Predicted Exons:  
INTR 51184 ... 50703 OCKHAMG-CDNA
- 10  
SINGLE 51148 ... 50810 OCKHAMG-CDS  
gi No: 5809708  
Predicted Exons:  
INTR 41416 ... 40935 OCKHAMG-CDNA
- 15  
SINGLE 41380 ... 41042 OCKHAMG-CDS  
gi No: 5836114  
Predicted Exons:  
INTR 42487 ... 42006 OCKHAMG-CDNA
- 20  
SINGLE 42451 ... 42113 OCKHAMG-CDS  
gi No: 5923662  
Predicted Exons:  
INTR 43106 ... 43587 OCKHAMG-CDNA
- 25  
SINGLE 43144 ... 43482 OCKHAMG-CDS  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 595  
- Ceres seq\_id 1026648  
- Alternative transcription start site(s) located in SEQ ID NO 595:  
-74,-72,-48,-42,76
- 30  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 596  
- Ceres seq\_id 1026649  
- Location of start within SEQ ID NO 595: at 39 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 596: at 59 aa.
- 35  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 597  
- Ceres seq\_id 1026650  
- Location of start within SEQ ID NO 595: at 78 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 597: at 46 aa.
- 40  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 598  
- Ceres seq\_id 1026651  
- Location of start within SEQ ID NO 595: at 81 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 598: at 45 aa.

- 307  
(Ba) Polypeptide Activities: Arabidopsis specific  
gene, dicot  
specific gene, plant
- 5  
specific gene.  
Maximum Length Sequence:  
related to:  
Clone IDs:  
18274  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 599  
- Ceres seq\_id 1027881  
- Alternative transcription start site(s) located in SEQ ID NO 599:  
-8,-4,2
- 10  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 600  
- Ceres seq\_id 1027882  
- Location of start within SEQ ID NO 599: at 3 nt.
- 15  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 601  
- Ceres seq\_id 1027883  
- Location of start within SEQ ID NO 599: at 36 nt.
- 20  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 602  
- Ceres seq\_id 1027884  
- Location of start within SEQ ID NO 599: at 105 nt.  
- Location of Signal Peptide Cleavage Site within SEQ ID NO 602: at 18 aa.
- 25  
(Ba) Polypeptide Activities: Arabidopsis specific  
gene, dicot  
specific gene, plant
- 30  
specific gene.  
Maximum Length Sequence:  
related to:  
Clone IDs:  
121894  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 603  
- Ceres seq\_id 1381797  
- Alternative transcription start site(s) located in SEQ ID NO 603:  
-41,18,19,22,26
- 35  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 604
- 40  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 604
- 45  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 604
- 50  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 604

308

- Ceres seq\_id 1381798
- Location of start within SEQ ID NO 603: at 70 nt.

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 605
- Ceres seq\_id 1381799
- Location of start within SEQ ID NO 603: at 82 nt.

5

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 606
- Ceres seq\_id 1381800
- Location of start within SEQ ID NO 603: at 109 nt.

10

(Ba) Polypeptide Activities: Arabidopsis specific  
gene, dicot

specific gene, plant

specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

31667

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 607
- Ceres seq\_id 1442747

25

- Alternative transcription start site(s) located in SEQ

ID NO 607:

3,5,6,11,62,64,65,66,69,70,71,72,74

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 608
- Ceres seq\_id 1442748
- Location of start within SEQ ID NO 607: at 2 nt.

30

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 609
- Ceres seq\_id 1442749
- Location of start within SEQ ID NO 607: at 140 nt.

35

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 610
- Ceres seq\_id 1442750
- Location of start within SEQ ID NO 607: at 224 nt.

40

(Ba) Polypeptide Activities: Arabidopsis specific  
gene, dicot

specific gene, plant

specific gene.

Maximum Length Sequence:

related to:

50

309

Clone IDs:

270354

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 611
- Ceres seq\_id 1459199

5

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 612
- Ceres seq\_id 1459200
- Location of start within SEQ ID NO 611: at 2 nt.

10

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 613
- Ceres seq\_id 1459201
- Location of start within SEQ ID NO 611: at 56 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 613: at 20 aa.

15

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 614
- Ceres seq\_id 1459202
- Location of start within SEQ ID NO 611: at 80 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 614: at 17 aa.

20

25

(Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot

specific gene, plant

specific gene.

30

Maximum Length Sequence:

related to:

Clone IDs:

12250

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 615
- Ceres seq\_id 1565605
- Alternative transcription start site(s) located in SEQ ID NO 615:  
-17

35

40

-17

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 616
- Ceres seq\_id 1565606
- Location of start within SEQ ID NO 615: at 133 nt.

45

## (B) Polypeptide Sequence

50

310

- Pat. Appln. SEQ ID NO 617
- Ceres seq\_id 1565607
- Location of start within SEQ ID NO 615: at 181 nt.

5 (Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot specific gene, plant

10 Maximum Length Sequence:

related to:

Clone IDs:

97883

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 618

- Ceres seq\_id 1566686

- Alternative transcription start site(s) located in SEQ

ID NO 618:

58, 61

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 619

- Ceres seq\_id 1566687

- Location of start within SEQ ID NO 618: at 137 nt.

25 (Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot specific gene, plant

30 Maximum Length Sequence:

related to:

Clone IDs:

12408

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 620

- Ceres seq\_id 1567367

- Alternative transcription start site(s) located in SEQ

ID NO 620:

-39, 6, 11, 14, 17, 25, 32, 33, 34, 35, 60

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 621

- Ceres seq\_id 1567368

- Location of start within SEQ ID NO 620: at 1 nt.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 622

- Ceres seq\_id 1567369

- Location of start within SEQ ID NO 620: at 2 nt.

311

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 623

- Ceres seq\_id 1567370

- Location of start within SEQ ID NO 620: at 65 nt.

5 (Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot specific gene, plant

10 Maximum Length Sequence:

related to:

Clone IDs:

19528

Public Genomic DNA:

gi No: 6091770

Predicted Exons:

INIT 21323 ...

INTR 21467 ...

INTR 21995 ...

TERM 22386 ...

21390

21564

22027

22468

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS

OCKHAMG-CDS



312

(Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot

specific gene.

5

Maximum Length Sequence:

related to:

Clone IDs:

42399

10 (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 626

- Ceres seq\_id 1571051

15 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 627

- Ceres seq\_id 1571052

- Location of start within SEQ ID NO 626: at 1 nt.

20 (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 628

- Ceres seq\_id 1571053

- Location of start within SEQ ID NO 626: at 16 nt.

- Location of Signal Peptide Cleavage Site within SEQ

25 ID NO 101: at 30 aa.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 629

- Ceres seq\_id 1571054

- Location of start within SEQ ID NO 626: at 139 nt.

30

(Ba) Polypeptide Activities: Arabidopsis specific

gene, dicot

specific gene, plant

35 specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

6487

40 (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 630

- Ceres seq\_id 1571100

- Alternative transcription start site(s) located in SEQ

ID NO 630:

5, 6

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 631

- Ceres seq\_id 1571101

- Location of start within SEQ ID NO 630: at 1 nt.

50

313

- Location of Signal Peptide Cleavage Site within SEQ  
ID NO 631: at 19 aa.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 632

- Ceres seq\_id 1571102

- Location of start within SEQ ID NO 630: at 82 nt.

5

(Ba) Polypeptide Activities: Arabidopsis specific  
gene, dicot

specific gene, plant

10

specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

39977

14890

34623

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 633

- Ceres seq\_id 1665272

- Alternative transcription start site(s) located in SEQ

ID NO 633:

25 3,15,32,33,34,35,37,38,39,40,41,42,43,44,45,46,47,48,49,50

51,60,61,63,66,69,70,71,72,74,75,76,78,79,84,85

- Clone 39977 starts at 35 and ends at 636 in cDNA.

- Clone 14890 starts at 43 and ends at in cDNA.

- Clone 34623 starts at 1 and ends at 598 in cDNA.

30

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 634

- Ceres seq\_id 1665273

- Location of start within SEQ ID NO 633: at 90 nt.

35

(Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot

specific gene, plant

specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

24562

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 635

- Ceres seq\_id 1713895

- Alternative transcription start site(s) located in SEQ

ID NO 635:

-7,2,3,5,7,8,13,17,29,30,32,34

50

314

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 636  
 - Ceres seq\_id 1713896  
 - Location of start within SEQ ID NO 635: at 3 nt.

5

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 637  
 - Ceres seq\_id 1713897  
 - Location of start within SEQ ID NO 635: at 57 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 637: at 20 aa.

10

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 638  
 - Ceres seq\_id 1713898  
 - Location of start within SEQ ID NO 635: at 81 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 638: at 17 aa.

15

- (Ba) Polypeptide Activities: Arabidopsis specific gene,  
 dicot  
 specific gene.

25 Maximum Length Sequence:

related to:  
 Clone IDs:

41879

- (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 639  
 - Ceres seq\_id 1923752

30

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 640  
 - Ceres seq\_id 1923753  
 - Location of start within SEQ ID NO 639: at 2 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 640: at 19 aa.

35

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 641  
 - Ceres seq\_id 1923754  
 - Location of start within SEQ ID NO 639: at 292 nt.

40

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 642  
 - Ceres seq\_id 1923755  
 - Location of start within SEQ ID NO 639: at 384 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 642: at 27 aa.

50

315

- (Ba) Polypeptide Activities: Arabidopsis specific gene,  
 dicot  
 specific gene.

5

Maximum Length Sequence:  
 related to:  
 Clone IDs:

150069

- (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 643  
 - Ceres seq\_id 1976816  
 - Alternative transcription start site(s) located in SEQ ID NO 643:  
 20,30,32

15

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 644  
 - Ceres seq\_id 1976817  
 - Location of start within SEQ ID NO 643: at 2 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 644: at 61 aa.

20

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 645  
 - Ceres seq\_id 1976818  
 - Location of start within SEQ ID NO 643: at 3 nt.

25

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 646  
 - Ceres seq\_id 1976819  
 - Location of start within SEQ ID NO 643: at 52 nt.

30

- (Ba) Polypeptide Activities: Arabidopsis specific gene,  
 dicot  
 specific gene.

35

Maximum Length Sequence:  
 related to:  
 Clone IDs:

109997

Public Genomic DNA:

gi No: 4263813

Predicted Exons:

INIT 82737 ... 82740 OCKHAMG-CDS  
 TERM 83475 ... 83623 OCKHAMG-CDS

- (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 647  
 - Ceres seq\_id 2025128

50

316

- Alternative transcription start site(s) located in SEQ ID NO 647:

-5,5,61,88

- Clone 109997 starts at 5 and ends at 566 in cDNA.

5

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 648

- Ceres seq\_id 2025129

- Location of start within SEQ ID NO 647: at 1 nt.

10

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 649

- Ceres seq\_id 2025130

- Location of start within SEQ ID NO 647: at 3 nt.

- Location of Signal Peptide Cleavage Site within SEQ

15

ID NO 649: at 21 aa.

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 650

- Ceres seq\_id 2025131

- Location of start within SEQ ID NO 647: at 316 nt.

20

(Ba) Polypeptide Activities: Arabidopsis specific

gene, dicot

specific gene, plant

25

specific gene.

Maximum Length Sequence:

related to:

Clone IDs:

41792

Public Genomic DNA:

gi No: 5762549

Predicted Exons:

INTR 41272 ...

41749 OCKHAMG-CDNA

gi No: 5776585

Predicted Exons:

INTR 55175 ...

55652 OCKHAMG-CDNA

gi No: 5801669

Predicted Exons:

INTR 97802 ...

98279 OCKHAMG-CDNA

gi No: 5809709

Predicted Exons:

INTR 97803 ...

98280 OCKHAMG-CDNA

gi No: 5932531

Predicted Exons:

INTR 90652 ...

91129 OCKHAMG-CDNA

gi No: 6013612

Predicted Exons:

INTR 9143 ...

9620 OCKHAMG-CDNA

gi No: 6016671

SUBSTITUTE SHEET (RULE 26)

317

Predicted Exons:

INTR 9266 ...

gi No: 6041828

Predicted Exons:

INTR 9266 ...

gi No: 6175159

Predicted Exons:

INTR 92644 ...

gi No: 6175159

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 651

- Ceres seq\_id 2025402

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 652

- Ceres seq\_id 2025403

- Location of start within SEQ ID NO 651: at 186 nt.

15

(B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 653

- Ceres seq\_id 2025404

- Location of start within SEQ ID NO 651: at 2 nt.

20

(Ba) Polypeptide Activities: Similar to ribose 5-phosphate isomerase

Maximum Length Sequence:

related to:

Clone IDs:

28572

Public Genomic DNA:

gi No: 4725940

Predicted Exons:

INTR 99197 ...

99167 OCKHAMG-CDS

INTR 99060 ...

98960 OCKHAMG-CDS

INTR 98863 ...

98753 OCKHAMG-CDS

INTR 98654 ...

97823 OCKHAMG-CDS

INTR 97730 ...

97477 OCKHAMG-CDS

INTR 96772 ...

96683 OCKHAMG-CDS

TERM 96157 ...

96086 OCKHAMG-CDS

gi No: 5823567

Predicted Exons:

INTR 99197 ...

99167 OCKHAMG-CDS

INTR 99060 ...

98960 OCKHAMG-CDS

INTR 98863 ...

98753 OCKHAMG-CDS

INTR 98654 ...

97823 OCKHAMG-CDS

INTR 97730 ...

97477 OCKHAMG-CDS

INTR 96772 ...

96683 OCKHAMG-CDS

TERM 96157 ...

96086 OCKHAMG-CDS

(Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 654

- Ceres seq\_id 2025479

50

SUBSTITUTE SHEET (RULE 26)

- 318  
- Alternative transcription start site(s) located in SEQ ID NO 654:  
1419  
- Clone 28572 starts at 1419 and ends at 1955 in cDNA.
- 5  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 655  
- Ceres seq\_id 2025480  
- Location of start within SEQ ID NO 654: at 1 nt.
- 10  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 656  
- Ceres seq\_id 2025481  
- Location of start within SEQ ID NO 654: at 226 nt.
- 15  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 657  
- Ceres seq\_id 2025482  
- Location of start within SEQ ID NO 654: at 271 nt.
- 20  
(Ba) Polypeptide Activities: Similar Cell Wall Plasma Membrane disconnecting CLCT Protein activities, and proline rich protein activities.
- 25  
Maximum Length Sequence:  
Public Genomic DNA:  
gi No: 6434247  
Predicted Exons:  
INTR 46663 ... 46055 OCKHAMG-CDNA
- 30  
SINGLE 46579 ... 46103 OCKHAMG-CDS
- 35  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 658  
- Ceres seq\_id 2032963
- 40  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 659  
- Ceres seq\_id 2032964  
- Location of start within SEQ ID NO 658: at 85 nt.
- 45  
(Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot  
specific gene, plant
- 50  
Maximum Length Sequence:  
Public Genomic DNA:  
gi No: 6449509  
Predicted Exons:

- 319  
INIT 32110 ... 32118 OCKHAMG-CDS  
INTR 32363 ... 32538 OCKHAMG-CDS  
INTR 33006 ... 33084 OCKHAMG-CDS  
TERM 33182 ... 33247 OCKHAMG-CDS  
gi No: 6453848  
Predicted Exons:  
INIT 34795 ... 34803 OCKHAMG-CDS  
INTR 35048 ... 35223 OCKHAMG-CDS  
INTR 35691 ... 35769 OCKHAMG-CDS  
TERM 35867 ... 35932 OCKHAMG-CDS  
gi No: 6456153  
Predicted Exons:  
INIT 32305 ... 32313 OCKHAMG-CDS  
INTR 32558 ... 32733 OCKHAMG-CDS  
INTR 33201 ... 33279 OCKHAMG-CDS  
TERM 33377 ... 33442 OCKHAMG-CDS
- 20  
(Ac) cDNA Polynucleotide Sequence  
- Pat. Appln. SEQ ID NO 660  
- Ceres seq\_id 2033706
- 25  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 661  
- Ceres seq\_id 2033707  
- Location of start within SEQ ID NO 660: at 1 nt.
- 30  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 662  
- Ceres seq\_id 2033708  
- Location of start within SEQ ID NO 660: at 34 nt.
- 35  
(B) Polypeptide Sequence  
- Pat. Appln. SEQ ID NO 663  
- Ceres seq\_id 2033709  
- Location of start within SEQ ID NO 660: at 160 nt.
- 40  
(Ba) Polypeptide Activities: Arabidopsis specific gene,  
dicot  
specific gene, plant
- 45  
Maximum Length Sequence:  
Public Genomic DNA:  
gi No: 4512646  
Predicted Exons:  
INTR 43130 ... 43387 OCKHAMG-CDNA  
INTR 43479 ... 43638 OCKHAMG-CDNA
- 50  
INIT 43224 ... 43387 OCKHAMG-CDS  
TERM 43479 ... 43494 OCKHAMG-CDS

- 320
- (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 664  
 - Ceres seq\_id 2043118
- 5 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 665  
 - Ceres seq\_id 2043119  
 - Location of start within SEQ ID NO 664: at 65 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 665: at 22 aa.
- 10 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 666  
 - Ceres seq\_id 2043120  
 - Location of start within SEQ ID NO 664: at 1 nt.
- 15 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 667  
 - Ceres seq\_id 2043121  
 - Location of start within SEQ ID NO 664: at 95 nt.
- 20 (Ba) Polypeptide Activities: Arabidopsis specific gene,  
 dicot  
 specific gene, plant
- 25 Maximum Length Sequence:  
 Public Genomic DNA:  
 gi No: 3985934  
 Predicted Exons:  
 INTR 40094 ... 40516 OCKHAMG-CDNA  
 INTR 40861 ... 41275 OCKHAMG-CDNA
- 30 INIT 40233 ... 40516 OCKHAMG-CDS  
 TERM 40861 ... 41077 OCKHAMG-CDS
- 35 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 668  
 - Ceres seq\_id 2047214
- 40 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 669  
 - Ceres seq\_id 2047215  
 - Location of start within SEQ ID NO 668: at 140 nt.
- 45 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 670  
 - Ceres seq\_id 2047216  
 - Location of start within SEQ ID NO 668: at 224 nt.
- 50 (B) Polypeptide Sequence

- 321
- Pat. Appln. SEQ ID NO 671  
 - Ceres seq\_id 2047217  
 - Location of start within SEQ ID NO 668: at 153 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 671: at 18 aa.
- 5 (Ba) Polypeptide Activities: Arabidopsis specific gene,  
 dicot  
 specific gene, plant
- 10 Maximum Length Sequence:  
 Public Genomic DNA:  
 gi No: 3985958  
 Predicted Exons:  
 INTR 8598 ... 8696 OCKHAMG-CDNA  
 INTR 8960 ... 9035 OCKHAMG-CDNA  
 INTR 9202 ... 9513 OCKHAMG-CDNA  
 INTR 9628 ... 9669 OCKHAMG-CDNA
- 20 INIT 8997 ... 9035 OCKHAMG-CDS  
 INTR 9202 ... 9513 OCKHAMG-CDS  
 TERM 9628 ... 9822 OCKHAMG-CDS
- 25 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 672  
 - Ceres seq\_id 2047438
- 30 (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 673  
 - Ceres seq\_id 2047439  
 - Location of start within SEQ ID NO 672: at 137 nt.
- 35 Maximum Length Sequence:  
 Public Genomic DNA:  
 gi No: 4757403  
 Predicted Exons:  
 INTR 35281 ... 34781 OCKHAMG-CDNA
- 40 SINGLE 35235 ... 35005 OCKHAMG-CDS
- 45 (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 674  
 - Ceres seq\_id 2049056
- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 675  
 - Ceres seq\_id 2049057  
 - Location of start within SEQ ID NO 674: at 47 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 675: at 22 aa.

322

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 149
- Ceres seq\_id 2049058
- Location of start within SEQ ID NO 674: at 59 nt.
- Location of Signal Peptide Cleavage Site within SEQ ID NO 676: at 18 aa.

5

## Maximum Length Sequence:

## Public Genomic DNA:

gi No: 4538972

## Predicted Exons:

INTR	32316 ...	32561	OCKHAMG-CDNA
INTR	32643 ...	32715	OCKHAMG-CDNA
INTR	33489 ...	33697	OCKHAMG-CDNA
INTR	33789 ...	33935	OCKHAMG-CDNA
INTR	34034 ...	34365	OCKHAMG-CDNA
INIT	32328 ...	32561	OCKHAMG-CDS
INTR	32643 ...	32715	OCKHAMG-CDS
INTR	33489 ...	33697	OCKHAMG-CDS
INTR	33789 ...	33935	OCKHAMG-CDS
TERM	34034 ...	34159	OCKHAMG-CDS

20

## (Ac) cDNA Polynucleotide Sequence:

- Pat. Appln. SEQ ID NO 677
  - Ceres seq\_id 2050386
- (B) Polypeptide Sequence
- Pat. Appln. SEQ ID NO 678
  - Ceres seq\_id 2050387
  - Location of start within SEQ ID NO 677: at 1 nt.

30

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 679
- Ceres seq\_id 2050388
- Location of start within SEQ ID NO 677: at 13 nt.

35

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 680
- Ceres seq\_id 2050389
- Location of start within SEQ ID NO 677: at 151 nt.

40

## (Ba) Polypeptide Activities: Arabidopsis specific gene,

dicot

specific gene, plant

specific gene.

Maximum Length Sequence:

Public Genomic DNA:

gi No: 4878038

50

323

## Predicted Exons:

INTR	22640 ...	22509	OCKHAMG-CDNA
INTR	22412 ...	22340	OCKHAMG-CDNA
INTR	22122 ...	21728	OCKHAMG-CDNA
INTR	22674 ...	22509	OCKHAMG-CDNA
INTR	22412 ...	22340	OCKHAMG-CDNA
INTR	22122 ...	21766	OCKHAMG-CDNA
INIT	22587 ...	22509	OCKHAMG-CDS
INTR	22412 ...	22340	OCKHAMG-CDS
TERM	22122 ...	21852	OCKHAMG-CDS

gi No: 6143825

## Predicted Exons:

INTR	22639 ...	22508	OCKHAMG-CDNA
INTR	22411 ...	22339	OCKHAMG-CDNA
INTR	22121 ...	21727	OCKHAMG-CDNA
INTR	22673 ...	22508	OCKHAMG-CDNA
INTR	22411 ...	22339	OCKHAMG-CDNA
INTR	22121 ...	21765	OCKHAMG-CDNA
INIT	22586 ...	22508	OCKHAMG-CDS
INTR	22411 ...	22339	OCKHAMG-CDS
TERM	22121 ...	21851	OCKHAMG-CDS

25

## (Ac) cDNA Polynucleotide Sequence

- Pat. Appln. SEQ ID NO 681
- Ceres seq\_id 2053353

30

## (B) Polypeptide Sequence

- Pat. Appln. SEQ ID NO 682
- Ceres seq\_id 2053354
- Location of start within SEQ ID NO 681: at 88 nt.

35

## (Ba) Polypeptide Activities: Arabidopsis specific gene,

dicot

specific gene, plant

specific gene.

40

## Maximum Length Sequence:

Public Genomic DNA:

gi No: 4115352

## Predicted Exons:

INIT	29289 ...	29287	OCKHAMG-CDS
INTR	27681 ...	27485	OCKHAMG-CDS
INTR	27312 ...	27193	OCKHAMG-CDS
INTR	26979 ...	26916	OCKHAMG-CDS
TERM	26654 ...	26646	OCKHAMG-CDS

50

324

- (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 683  
 - Ceres seq\_id 2055693

5

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 684  
 - Ceres seq\_id 2055694  
 - Location of start within SEQ ID NO 683: at 1 nt.

10

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 685  
 - Ceres seq\_id 2055695  
 - Location of start within SEQ ID NO 683: at 46 nt.

15

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 686  
 - Ceres seq\_id 2055696  
 - Location of start within SEQ ID NO 683: at 2 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 686: at 28 aa.

20

(Ba) Polypeptide Activities: Arabidopsis specific gene, dicot

25

specific gene, plant

Maximum Length Sequence:

Public Genomic DNA:

gi No: 4406790

Predicted Exons:

INTR 81640 ... 81247 OCKHAMG-CDNA

SINGLE 81603 ... 81367 OCKHAMG-CDS

35

- (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 687  
 - Ceres seq\_id 2056405

40

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 688  
 - Ceres seq\_id 2056406  
 - Location of start within SEQ ID NO 687: at 38 nt.

45

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 689  
 - Ceres seq\_id 2056407  
 - Location of start within SEQ ID NO 687: at 56 nt.

50

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 690

325

- Ceres seq\_id 2056408  
 - Location of start within SEQ ID NO 687: at 177 nt.  
 - Location of Signal Peptide Cleavage Site within SEQ ID NO 690: at 14 aa.

5

- (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot  
 specific gene, plant

10

Maximum Length Sequence:

related to:

Clone IDs:

233233

15

- (Ac) cDNA Polynucleotide Sequence  
 - Pat. Appln. SEQ ID NO 691  
 - Ceres seq\_id 2065747

20

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 692  
 - Ceres seq\_id 2065748  
 - Location of start within SEQ ID NO 691: at 114 nt.

25

- (B) Polypeptide Sequence  
 - Pat. Appln. SEQ ID NO 693  
 - Ceres seq\_id 2065749  
 - Location of start within SEQ ID NO 691: at 279 nt.

30

- (Ba) Polypeptide Activities: Arabidopsis specific gene, dicot  
 specific gene, plant

326

## CLAIMS

What is claimed is:

1. An isolated nucleic acid molecule comprising a nucleic acid having a nucleotide sequence selected from the group consisting of
  - (i) a nucleotide sequence shown in the Sequence Listing;
  - (ii) a nucleotide sequence complementary to one shown in the Sequence Listing;
  - (iii) a nucleotide sequence which is the reverse complement of one shown in the Sequence Listing;
  - (iv) a nucleotide sequence capable of hybridizing to a nucleotide sequence shown in the Sequence Listing;
  - (iv) a nucleotide sequence capable of hybridizing to a nucleotide sequence complementary to one shown in the Sequence Listing;
  - (vi) a nucleotide sequence capable of hybridizing to a nucleotide sequence which is the reverse complement of one shown in the Sequence Listing;
 whereby the hybridization in groups (iv) to (vi) allow said sequences to form a duplex at a temperature from about T<sub>m</sub>-40°C to about T<sub>m</sub>-48°C.
2. An isolated nucleic acid molecule according to claim 1 that comprises a complete open reading frame.
3. An isolated nucleic acid molecule according to claim 1 that functions as a promoter or as a 3' end termination sequence or as a regulator sequence influencing the transcription rate, the transcript stability or RNA translation rate in a host cell.

SUBSTITUTE SHEET (RULE 26)

327

4. The isolated nucleic acid of claim 3, comprising a sequence selected from the group consisting of a TATA box sequence, a CAAT box sequence, a motif of GCAATCG or a transcription-factor binding sequence.
5. An isolated nucleic acid molecule comprising a nucleic acid sequence that encodes an amino acid sequence selected from the group consisting of the amino acid sequences shown in the Sequence Listing, a complementary nucleotide sequence to said nucleic acid sequence that encodes said amino acid sequence selected from the Sequence Listing, a reverse complementary nucleotide sequence to said nucleic acid sequence that encodes said amino acid sequence selected from the Sequence Listing, and a nucleotide sequence that will hybridize to said nucleic acid sequence encoding an amino acid selected from the Sequence Listing or said complementary sequence under hybridization conditions providing T<sub>m</sub> - 40°C to T<sub>m</sub> - 48°C.
6. An isolated nucleic acid molecule comprising a nucleic acid sequence that encodes at least a part or a portion or a mutant or a fusion of an amino acid sequence selected from the group consisting of the amino acid sequences shown in the Sequence Listing, a complementary nucleotide sequence to said nucleic acid sequence that encodes said part or portion or mutant or fusion of an amino acid sequence selected from the Sequence Listing, a reverse complementary nucleotide sequence to said nucleic acid sequence that encodes said part or portion or mutant or fusion of an amino acid sequence selected from the Sequence Listing, and a nucleotide sequence that will hybridize to said nucleic acid sequence that encodes said part or portion or mutant or fusion of an amino acid selected from the Sequence Listing or said complementary sequence or said reverse complementary sequence under hybridization conditions providing T<sub>m</sub> - 40°C to T<sub>m</sub> - 48°C.

SUBSTITUTE SHEET (RULE 26)



328

7. An isolated nucleic acid molecule according to claim 1, having a nucleotide sequence selected from the group consisting of the nucleotide sequences shown in the Sequence Listing, and a complementary nucleotide sequence to said nucleotide sequence selected from the Sequence Listing.

8. A chimeric DNA construct comprising (a) a promoter sequence capable of driving gene expression in plant cells and operatively linked to (b) a structural gene comprising an DNA molecule according to any one of claims 1, 2, 5 and 7.

9. A chimeric DNA construct comprising (a) a promoter that is a nucleic acid molecule according to claim 3 or 4 operatively linked to (b) a structural gene and, optionally, (c) an associated UTR.

10. An isolated DNA molecule comprising (a) a promoter sequence according to claim 3 or 4 and operatively linked to (b) a structural gene sequence according to any one of claims 1, 2, 5, 6 and 7.

11. A promoter according to claim 10, wherein said promoter is a specific promoter.

12. A promoter according to claim 11, wherein said promoter is a seed-specific promoter, an embryo-specific promoter, an ovule-specific promoter, a tapetum-specific promoter or a root-specific promoter.

13. A host cell transformed with a polynucleotide comprising the isolated nucleic acid molecule according to claim 1.

329

14. An isolated polypeptide comprising an amino acid sequence encoded by a polynucleotide sequence shown in the Sequence Listing, or one that is at least 75% identical thereto.

15. An isolated polypeptide according to claim 14, that is at least 85% identical to said amino acid sequence.

16. An isolated polypeptide, according to claim 15, that is at least 90% identical to said amino acid sequence.

17. A polynucleotide comprising a first polynucleotide sequence from the Sequence Listing or a fragment thereof, wherein said first sequence is capable of regulating transcription or translation of second a polynucleotide comprising a coding sequence in a host cell or in vitro system wherein the first and second polynucleotides are operatively linked.

18. The first polynucleotide of claim 17, which is capable of regulating transcription or translation in an in vitro system.

19. The first polynucleotide of claim 17, which is capable of regulating transcription or translation in a host cell.

20. The polynucleotide of claim 19, wherein said host cell is a plant cell.

21. A host cell of claim 13, wherein said isolated nucleic acid comprises a promoter and operatively linked structural gene and further wherein said promoter and

330

structural gene are not heterologous to each other and are  
5 exogenous to the host cell genome.

331

28. The isolated polypeptide according to claim 26 that  
is at least 90% identical thereto.

22. A method of introducing an isolated nucleic acid into  
a host cell comprising:

(a) providing an isolated nucleic acid of any of claims  
1-12;

5 (b) contacting said isolated nucleic acid with said host  
cell under conditions that permit insertion of said nucleic  
acid into said host cell.

23. A method of producing a polypeptide of any one of  
claims 14-16 comprising:

(a) providing a host cell of claim 13;

5 (b) culturing said host cell under conditions that permit  
transcription and translation of said structural gene to  
produce a polypeptide; and

(c) isolating said polypeptide.

24. An antibody capable of binding to a polypeptide of  
any one of claims 14-16.

25. An isolated nucleic acid comprising a promoter of a  
gene wherein said gene comprises a nucleic acid having the  
nucleotide sequence of claim 1.

26. An isolated polypeptide comprising an amino acid  
sequence selected from the Sequence Listing, or one that is at  
least 75% identical thereto.

27. The isolated polypeptide according to Claim 26 that  
is at least 85% identical thereto.

<110> ALEXANDROV, Nikolai  
 BROVER, Vyacheslav  
 CHEN, Xianfeng  
 SUBRAMANIAN, Gopalakrishnan  
 TROUKHAN, Maxim  
 ZHENG, Liansheng

Leu Gly Gln Val Leu Met Arg Ala Arg Val Val Gln Gly Ala Thr  
 90 95  
 Val Ala Leu Met Val Gly Thr Gly Tyr Tyr Gly Asp Asn Pro Trp  
 100 105  
 Lys Lys

<120> SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING  
 POLYPEPTIDES ENCODED THEREBY

<130> 2750-686F(PC)

<140> PCT/US00/00466

<141> 2000-01-07

<160> 693

<210> 1

<211> 601

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..601:Ceres Seq. ID 1007546

<400>1

gaagacagg ccacgcgtga ctacatcat aatttggtt cttcttttt ttgtgtctt  
 60  
 cttgtctc ttgtgtttt gaattggtt ctggagaaga tgaattcgtt agaaccagac  
 120  
 atggaagac tgtccagga gaagaagct gtccagaatc ctccgttcc tctgttga  
 180  
 cttatgacc cggagtggt caggtctgg ttgattagt tcagaagagg caattctcag  
 240  
 ttgttcagg ttgtatgag agctagggtg gtcgtccagg gtgtactgt cgttttaag  
 300  
 gtggaaccy gttattacda cgtgtataat cgttgaaga agtgagctcc aactattgt  
 360  
 tegtttcac ataattgaga tgacactcgg gaactcggc ggcagaagtg tgtgtctct  
 420  
 cgtctcagt gtagaatat gctgtctgc taggatttag taagggtaca ttacatgatt  
 480  
 tgaataatt gaagaacgaa caagtatcgt tgtatctga tcaattgatt agattgatt  
 540  
 aaggaaattc tcggaanacc atctttactc tatctgtcaa aattctatta tcttttgc  
 600  
 g

<210> 2

<211> 114

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..114:Ceres Seq. ID 1007547

<400>2

Asp Glu Gln Ala Thr Arg Asp Tyr Ile His Asn Leu Val Ser Leu Phe  
 1 5 10 15  
 Phe Phe Val Leu Leu Val Pro Leu Trp Phe Leu Thr Leu Val Leu Glu  
 20 25 30  
 Lys Met Ser Val Glu Pro Asp Met Glu Asp Leu Phe Gln Glu Lys  
 35 40 45  
 Lys Arg Val Arg Asn Pro Leu Val Pro Leu Gly Ala Leu Met Thr Ala  
 50 55 60  
 Gly Val Leu Thr Ala Gly Ile Ser Phe Arg Arg Gly Asn Ser Gln  
 65 70 75 80

<210> 3

<211> 81

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..81:Ceres Seq. ID 1007548

<400>3

Met Ser Ser Val Glu Pro Asp Met Glu Asp Leu Phe Gln Glu Lys Lys  
 1 5 10 15  
 Arg Val Arg Asn Pro Leu Val Pro Leu Gly Ala Leu Met Thr Ala Gly  
 20 25 30  
 Val Leu Thr Ala Gly Leu Ile Ser Phe Arg Arg Gly Asn Ser Gln Leu  
 35 40 45  
 Gly Gln Val Leu Met Arg Ala Arg Val Val Gln Gly Ala Thr Val  
 50 55 60  
 Ala Leu Met Val Gly Thr Gly Tyr Tyr Gly Asp Asn Pro Trp Lys  
 65 70 75 80  
 Lys

<210> 4

<211> 74

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..74:Ceres Seq. ID 1007549

<400>4

Met Glu Asp Leu Phe Gln Glu Lys Lys Arg Val Arg Asn Pro Leu Val  
 1 5 10 15  
 Pro Leu Gly Ala Leu Met Thr Ala Gly Val Leu Thr Ala Gly Leu Ile  
 20 25 30  
 Ser Phe Arg Arg Gly Asn Ser Gln Leu Gly Gln Val Leu Met Arg Ala  
 35 40 45  
 Arg Val Val Val Gln Gly Ala Thr Val Ala Leu Met Val Gly Thr Gly  
 50 55 60  
 Tyr Tyr Tyr Gly Asp Asn Pro Trp Lys Lys  
 65 70

<210> 5

<211> 970

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..970:Ceres Seq. ID 1007583

<400>5  
 aacgcgcctc atttttaggg ttccagatat aggcagattc tccatttgc aaacatgacg 60  
 actcagata gcaagaagag aagttttaga gggcaggtg tttctacgc tgaattgaat 120  
 gaggttctta caagagagct agcagaggtt ggttactctg gtgttgaggt taggttact 180  
 cctatgagga ctgagattat catcagagct acacgtactc agaatgtct cgttgagaag 240  
 gggagggaaa ttagggaatt gacttccctt gtccagaaga gattccagtt tccagtgtac 300  
 agttgtgacg ttatagccga gaaggttaac aacagagtc tctgtgccat cgtccaggt 360  
 gagctttac gtacaaagct tctgttgtgt ctgcgtgttc gttaggcttg ctatgggttg 420  
 ttgaggttg ttatgagag ttgagctaa ggtatcgagg gtatcgtag tgaagaactt 480  
 cgtgtgccca gagccaagtc tatgaagttc aaagtggct acatgggtc actgggtcaa 540  
 ccaactaagg aatacataga cctgcagtg agacatgttt tgcctagaca aggtgtgtg 600  
 ggaatcaagg tgaagttat ccttgatctt cctcaagaag agaggtccat ttaagcact 720  
 cctatgcctg atgttgtgat cattcatctt gaccctaagg gcatatcagg accaaagaca 780  
 gctcaggttg ctgcctgcgc tgcctctgta gcatgtgcac cactcacagc cgtatgttac 840  
 cctcagatga tccagctgcg ctaaaagaat ccttgttact agtcagatgt atttgtttt 900  
 caactgtctt tgcctcttgg acagattata gcttctctga tcttcgaatc agacagttt 960  
 cctctgtctt tattccgaaa ttatctcaag ttctcattgt tgaaaaagcc ctttttagtt 970  
 tcaataatcg

&lt;210&gt; 6

&lt;211&gt; 249

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..249:Ceres Seq. ID 1007584

<400>6  
 Met Thr Thr Gln Ile Ser Lys Lys Arg Lys Phe Val Ala Asp Gly Val 1  
 1 5 10 15  
 Phe Tyr Ala Glu Leu Asn Glu Val Leu Thr Arg Glu Leu Ala Glu Asp 20  
 Gly Tyr Ser Gly Val Glu Val Arg Val Thr Pro Met Arg Thr Glu Ile 30  
 35 40 45  
 Ile Ile Arg Ala Thr Arg Thr Gln Asn Val Leu Gly Glu Lys Gly Arg 40  
 50 55  
 Arg Ile Arg Glu Leu Thr Ser Leu Val Gln Lys Arg Phe Lys Phe Pro 60  
 65 70 75 80  
 Val Asp Ser Val Glu Leu Tyr Ala Glu Lys Val Asn Asn Arg Gly Leu 85  
 90  
 Cys Ala Ile Ala Gln Ala Glu Ser Leu Arg Tyr Lys Leu Leu Gly Gly 100  
 105 110  
 Leu Ala Val Arg Arg Ala Cys Tyr Gly Val Leu Arg Phe Val Met Glu 115  
 120 125  
 Ser Gly Ala Lys Gly Cys Glu Val Ile Val Ser Gly Lys Leu Arg Ala 130  
 135 140  
 Ala Arg Ala Lys Ser Met Lys Phe Lys Asp Gly Tyr Met Val Ser Ser 145  
 150 155 160  
 Gly Gln Pro Thr Lys Glu Tyr Ile Asp Ser Ala Val Arg His Val Leu 165  
 170 175  
 Leu Arg Gln Gly Val Leu Gly Ile Lys Val Lys Val Met Leu Asp Trp 180  
 185 190  
 Asp Pro Lys Gly Ile Ser Gly Pro Lys Thr Pro Leu Pro Asp Val Val 195  
 200 205  
 Ile Ile His Ser Pro Lys Glu Glu Ala Ile Tyr Ala Pro Ala Gln 210  
 215 220  
 Val Ala Ala Pro Ala Leu Val Ala Asp Ala Pro Leu Thr Ala Val

&lt;223&gt; LOCATION: 1..206:Ceres Seq. ID 1007585

<400>7  
 Met Arg Thr Glu Ile Ile Arg Ala Thr Arg Thr Gln Asn Val Leu 1  
 1 5 10 15  
 Gly Glu Lys Gly Arg Arg Ile Arg Glu Leu Thr Ser Leu Val Gln Lys 20  
 25 30  
 Arg Phe Lys Phe Pro Val Asp Ser Val Glu Leu Tyr Ala Glu Lys Val 35  
 40 45  
 Asn Asn Arg Gly Leu Cys Ala Ile Ala Gln Ala Glu Ser Leu Arg Tyr 50  
 55 60  
 Lys Leu Leu Gly Gly Leu Ala Val Arg Arg Ala Cys Tyr Gly Val Leu 65  
 70 75 80  
 Arg Phe Val Met Glu Ser Gly Ala Lys Gly Cys Glu Val Ile Val Ser 85  
 90 95  
 Gly Lys Leu Arg Ala Arg Ala Lys Ser Met Lys Phe Lys Asp Gly 100  
 105 110  
 Tyr Met Val Ser Ser Gly Gln Pro Thr Lys Glu Tyr Ile Asp Ser Ala 115  
 120 125  
 Val Arg His Val Leu Leu Arg Gln Gly Val Leu Gly Ile Lys Val Lys 130  
 135 140  
 Val Met Leu Asp Trp Asp Pro Lys Gly Ile Ser Gly Pro Lys Thr Pro 145  
 150 155 160  
 Leu Pro Asp Val Val Ile Ile His Ser Pro Lys Glu Glu Ala Ile 165  
 170 175  
 Tyr Ala Pro Ala Gln Val Ala Ala Pro Ala Ala Leu Val Ala Asp Ala 180  
 185 190  
 Pro Leu Thr Ala Val Asp Tyr Pro Ala Met Ile Pro Val Ala 195  
 200 205  
 <210> 8  
 <211> 768  
 <212> DNA  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..768:Ceres Seq. ID 1008148  
 <400>8  
 aaaaaacc tagtttttg cttctcga gttccgtgcg actctctgtc cgaagaagt 60  
 ctctaccga tacaatgatt atctcagagg ctaaccgcaa agaatctcgc agtacctct 120  
 tcaagaagg agtttgcttt gcgaagaagg attcaatctt tctaagcat cegtgtatcg 180  
 atgtcccaa cctacaagtg attagaetta tgcagagttt caaatccaag gagtatgtta 240  
 gagagacatt tgcctggatg cattactatt ggittttgac caatgaagg atcgagttct 300  
 tgaagaacta tcttaattct caatatgat ttgttctgc tactttgaag aagtcagcta 360  
 agctctggcg tctccattt ggtggccac cgttgatcg ctaagaagg cctgcgcag 420  
 aaggaggaga cgtgccagg ttgttgacc gtgatggta cgtgtcaggt cctcagctg 480

gtggtgagtt tgaagtgaag aagggtgag ctcctgcaga ttaccagcca tctttccaag 540  
 gaagtgggg tggttttggc cgggtgtcg ggggtacag cgcagtgca ccatctggtt 600  
 caggtttgcc tgaanaaga aatgcttta ggtgacagta agaccatgga ggaagtttcca 660  
 gctttaatt tgcctttgt aattcagctt ccggaatcct tcataatctc tatctgagtt 720  
 tagttttgtt gttgaatcaa acatccgatt taagtttatg tteattcc 768

<210> 9  
 <211> 203  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..203:Ceres Seq. ID 1008149

<400>9  
 Lys Asn Pro Ser Phe Leu Leu Ser Arg Ser Ser Val Arg Leu Ser Ala 1  
 1  
 Arg Ala Arg Ser Pro Thr Asp Thr Met Ile Ile Ser Glu Ala Asn Arg 15  
 20  
 Lys Glu Ile Cys Lys Tyr Leu Phe Lys Glu Gly Val Cys Phe Ala Lys 30  
 35  
 Lys Asp Phe Asn Leu Ala Lys His Pro Leu Ile Asp Val Pro Asn Leu 45  
 50  
 Gln Val Ile Lys Leu Met Gln Ser Phe Lys Ser Lys Glu Tyr Val Arg 60  
 65  
 Glu Thr Phe Ala Trp Met His Tyr Tyr Trp Phe Leu Thr Asn Glu Gly 80  
 85  
 Ile Glu Phe Leu Arg Thr Tyr Leu Asn Leu Pro Ser Asp Val Val Pro 95  
 100  
 Ala Thr Leu Lys Lys Ser Ala Lys Pro Gly Gly Arg Pro Phe Gly Gly 110  
 115  
 Pro Pro Gly Asp Arg Ser Arg Gly Pro Arg His Glu Gly Gly Asp Arg 125  
 130  
 Pro Arg Phe Gly Asp Arg Asp Gly Tyr Arg Ala Gly Pro Arg Ala Gly 140  
 145  
 Gly Glu Phe Gly Gly Glu Lys Gly Gly Ala Pro Ala Asp Tyr Gln Pro 155  
 160  
 Ser Phe Gln Gly Ser Gly Arg Gly Phe Gly Arg Gly Ala Gly Gly Tyr 175  
 180  
 Ser Ala Ala Pro Ser Gly Ser Gly Leu Pro 190  
 195

<210> 10  
 <211> 179  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..179:Ceres Seq. ID 1008150

<400>10  
 Met Ile Ile Ser Glu Ala Asn Arg Lys Glu Ile Cys Lys Tyr Leu Phe 1  
 5  
 Lys Glu Gly Val Cys Phe Ala Lys Lys Asp Phe Asn Leu Ala Lys His 15  
 20  
 Pro Leu Ile Asp Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln Ser 30  
 35

Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His Tyr 50  
 Tyr Trp Phe Leu Thr Asn Glu Gly Ile Glu Phe Leu Arg Thr Tyr Leu 60  
 65  
 Asn Leu Pro Ser Asp Val Val Pro Ala Thr Leu Lys Lys Ser Ala Lys 75  
 80  
 Pro Gly Gly Arg Pro Phe Gly Gly Pro Pro Gly Asp Arg Ser Arg Gly 95  
 100  
 Pro Arg His Glu Gly Gly Asp Arg Pro Arg Phe Gly Asp Arg Asp Gly 110  
 115  
 Tyr Arg Ala Gly Pro Arg Ala Gly Gly Glu Phe Phe Gly Glu Lys Gly 125  
 130  
 Gly Ala Pro Ala Asp Tyr Gln Pro Ser Phe Gln Gly Ser Gly Arg Gly 140  
 145  
 Phe Gly Arg Gly Ala Gly Tyr Ser Ala Ala Pro Ser Gly Ser 155  
 160  
 Gly Leu Pro 170  
 175

<210> 11  
 <211> 134  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..134:Ceres Seq. ID 1008151

<400>11  
 Met Gln Ser Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp 1  
 5  
 Met His Tyr Tyr Trp Phe Leu Thr Asn Glu Gly Ile Glu Phe Leu Arg 15  
 20  
 Thr Tyr Leu Asn Leu Pro Ser Asp Val Val Pro Ala Thr Leu Lys Lys 25  
 35  
 Ser Ala Lys Pro Gly Gly Arg Pro Phe Gly Gly Pro Pro Gly Asp Arg 40  
 45  
 Ser Arg Gly Pro Arg His Glu Gly Gly Asp Arg Pro Arg Phe Gly Asp 50  
 55  
 65  
 Arg Asp Gly Tyr Arg Ala Gly Pro Arg Ala Gly Gly Glu Phe Gly Gly 75  
 80  
 Glu Lys Gly Gly Ala Pro Ala Asp Tyr Gln Pro Ser Phe Gln Gly Ser 95  
 100  
 Gly Arg Gly Phe Gly Arg Gly Ala Gly Gly Tyr Ser Ala Ala Pro 110  
 115  
 Ser Gly Ser Gly Leu Pro 120  
 125  
 130

<210> 12  
 <211> 995  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..995:Ceres Seq. ID 1008134

<400>12  
 atatctaaaa gactcaaaaa cctatcgta ttatatctca acagcgagaa gaacgggat 60

ggcgccatt acagctctca ctctcgctc tctgtttt ttctctctc catcgctcgt 120  
 cacaagccat agattccatg gcttcaacca tcaaccacca ccagctgctc tctcttcc 180  
 tctaaaccc ttctctctc tctcaatca aaacctataa tcaatccgaa ttctcgatc 240  
 cgcttgcg ataacacac caatctcca gagcgagaag tcaacagctc ggtcatcgac 300  
 actcaocgg tccactcgt ctctcgctac ctccgpgct ttgcaatcg ctgtaaccag 360  
 agtcttagt cagaactct ctctcgat caagcttca agtccgtaa tgcgcagcg 420  
 gtacagtc tctctagta ccgcggacc tgtctcttc ggcctctcc gggatcgcc 480  
 cccggatac ttgaacacg cggtgcggt tgtcggtg gggataaga agtgcctaga 540  
 taatcaatt cctctctat tcatatcgt tgaattagt ccaattggt ctctcgcggt 600  
 tcttgaca ctggatga ttgtcatgg cagcatgga tagaaatgg aagactgag 780  
 ctaatggt ttgtgaacg ttaattgg aaatttgc aacttggt coagattag 840  
 ggaataca gaattgag ctgtgtga tgcctgata ggtgtgag ctctatggt 900  
 ttgtgttc tctttgac ttgatttg aataatgaga ttgtggtg ttgttatct 960  
 cctaataca tgttcagact acagtatt gatgg 995

<210> 13

<211> 253

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..253:Ceres Seq. ID 1008335

<400>13

Tyr Leu Lys Asp Ser Lys Thr Tyr Arg His Leu Ser Ser Thr Gly Gly 15  
 1 5 10  
 Arg Thr Ala Met Ala Ala Ile Thr Ala Leu Thr Leu Arg Ser Pro Val 20  
 20 25 30  
 Phe Leu Leu Pro Pro Ser Ser Val Thr Ser Pro Arg Phe His Gly Phe 35  
 35 40 45  
 Thr Asn Gln Pro Pro Pro Ala Arg Leu Phe Phe Pro Leu Asn Pro Phe 50  
 50 55 60  
 Pro Ser Leu Ser Ile Gln Asn Pro Lys Ser Ile Arg Ile Ser Ala Ser 65  
 65 70 75  
 Ala Ser Pro Ile Thr Thr Pro Ile Leu Gln Thr Glu Lys Ser Thr Ala 80  
 85 90 95  
 Arg Ser Ser Thr Leu Thr Gly Ser Thr Arg Ser Leu Ala Thr Leu Ala 100  
 100 105 110  
 Ala Leu Ala Ile Ala Val Thr Arg Val Leu Ala Gln Lys Leu Ser Leu 115  
 115 120 125  
 Ala Ile Gln Thr Ser Ser Pro Val Ile Ala Asp Gly Leu Arg Phe Ser 130  
 130 135 140  
 Leu Ser Thr Ala Gly Pro Val Phe Phe Ala Ser Leu Arg Asp Arg Pro 145  
 145 150 155  
 Pro Gly Tyr Leu Asn Thr Pro Leu Thr Val Val Ala Val Gly Ile Lys 160  
 165 170 175  
 Lys Trp Leu Asp Ile Tyr Ser Gly Val Leu Met Val Arg Val Leu Leu 180  
 180 185 190  
 Ser Trp Phe Pro Asn Ile Pro Trp Glu Arg Gln Pro Leu Ser Ala Ile 195  
 195 200 205  
 Arg Asp Leu Cys Asp Pro Tyr Leu Asn Leu Phe Arg Asn Ile Ile Pro 210  
 210 215 220  
 Pro Ile Phe Asp Thr Leu Asp Val Ser Pro Leu Leu Ala Phe Ala Val 225  
 225 230 235  
 Leu Gly Thr Leu Gly Ser Ile Val His Gly Ser Thr Gly 240  
 245 250 255

<210> 14  
 <211> 234  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..234:Ceres Seq. ID 1008336  
 <400>14  
 Met Ala Ala Ile Thr Ala Leu Thr Leu Arg Ser Pro Val Phe Leu Leu 15  
 1 5 10  
 Pro Pro Ser Ser Val Thr Ser Pro Arg Phe His Gly Phe Thr Asn Gln 20  
 20 25 30  
 Pro Pro Pro Ala Arg Leu Phe Phe Pro Leu Asn Pro Phe Pro Ser Leu 35  
 35 40 45  
 Ser Ile Gln Asn Pro Lys Ser Ile Arg Ile Ser Ala Ser Ala Ser Pro 50  
 50 55 60  
 Ile Thr Thr Pro Ile Leu Gln Thr Glu Lys Ser Thr Ala Arg Ser Ser 65  
 65 70 75  
 Thr Leu Thr Gly Ser Thr Arg Ser Leu Ala Thr Leu Ala Ala Leu Ala 85  
 85 90 95  
 Ile Ala Val Thr Arg Val Leu Ala Gln Lys Leu Ser Leu Ala Ile Gln 100  
 100 105 110  
 Thr Ser Ser Pro Val Ile Ala Asp Gly Leu Arg Phe Ser Leu Ser Thr 115  
 115 120 125  
 Ala Gly Pro Val Phe Phe Ala Ser Leu Arg Asp Arg Pro Pro Gly Tyr 130  
 130 135 140  
 Leu Asn Thr Pro Leu Thr Val Ala Val Gly Ile Lys Lys Trp Leu 145  
 145 150 155  
 Asp Ile Tyr Ser Gly Val Leu Met Val Arg Val Leu Leu Ser Trp Phe 165  
 165 170 175  
 Pro Asn Ile Pro Trp Glu Arg Gln Pro Leu Ser Ala Ile Arg Asp Leu 180  
 180 185 190  
 Cys Asp Pro Tyr Leu Asn Leu Phe Arg Asn Ile Ile Pro Pro Ile Phe 195  
 195 200 205  
 Asp Thr Leu Asp Val Ser Pro Leu Leu Ala Phe Ala Val Leu Gly Thr 210  
 210 215 220  
 Leu Gly Ser Ile Val His Gly Ser Thr Gly 225  
 225 230 235

<210> 15

<211> 520

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..520:Ceres Seq. ID 1008701

<400>15

acacataacc ctcttgaaa gagtctaac acttgcagag aaaaagaaga aggaagatcc 60  
 cgaaatagg caacggcat tttacttca gctcttccc gagcagcag tcgcagct 120  
 ctgaagatc cgtctctcc taagcgaac ttctctctt ccgcgcgcca tgcagtgct 180  
 taagaagtg cgaagtggga gaagtaact tatctggta ttgtagtgt cactgtcta 240  
 gctgtctat ttatccaa gggccatcat ccgcgcgaag acctctgc ctatccgat 300  
 atcacatcc gcaacaaga gttctctgg ggtccgagtg gtcgtttga ggtgaagcac 360  
 aacaagac actgagttt ggtgtgcat aataactct tctgtgta ttgaagac 420  
 taaatggtt tccgtattt gttctcaacg ttgtcaacg atttctact ccaatctct 480

ttcttttgtt gggaaataaa agttaataact ttgttggtc

520

&lt;210&gt; 16

&lt;211&gt; 124

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..124:Ceres Seq. ID 1008702

&lt;400&gt;16

Thr His Asn Pro Leu Leu Gly Lys Ser Leu Asn Thr Cys Arg Glu Lys Glu

1 5 10 15

Glu Gly Arg Ser Arg Lys Met Ala Thr Ala Ile Val Arg Ser Ala Leu

20 25 30

Ser Arg Ala Ala Thr Arg Ala Ala Pro Lys Thr Ser Val Ala Pro Lys

35 40 45

Arg Asn Phe Ser Ser Ala Gly His Asp Asp Ala Tyr Glu Ala Ala

50 55 60

Lys Trp Glu Lys Ile Thr Tyr Leu Gly Ile Ala Ser Cys Thr Ala Leu

65 70 75 80

Ala Val Tyr Val Leu Ser Lys Gly His His Pro Gly Glu Asp Pro

85 90 95

Ala Tyr Pro His Met His Ile Arg Asn Lys Glu Phe Pro Trp Gly Pro

100 105 110

Asp Gly Leu Phe Glu Val Lys His Asn Lys Glu His

115 120

&lt;210&gt; 17

&lt;211&gt; 102

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..102:Ceres Seq. ID 1008703

&lt;400&gt;17

Met Ala Thr Ala Ile Val Arg Ser Ala Leu Ser Arg Ala Ala Thr Arg

1 5 10 15

Ala Ala Pro Lys Thr Ser Val Ala Pro Lys Arg Asn Phe Ser Ser

20 25 30

Ala Gly His Asp Asp Ala Tyr Glu Ala Ala Lys Trp Glu Lys Ile Thr

35 40 45

Tyr Leu Gly Ile Ala Ser Cys Thr Ala Leu Ala Val Tyr Val Leu Ser

50 55 60

Lys Gly His His Pro Gly Glu Asp Pro Ala Tyr Pro His Met His

65 70 75 80

Ile Arg Asn Lys Glu Phe Pro Trp Gly Pro Asp Gly Leu Phe Glu Val

85 90 95

Lys His Asn Lys Glu His

100

&lt;210&gt; 18

&lt;211&gt; 68

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

9

SUBSTITUTE SHEET (RULE 26)

WO 00/40695

<220>  
<223> any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..68:Ceres Seq. ID 1008704

&lt;400&gt;18

His Ile Thr Leu Leu Glu Arg Val Ser Thr Leu Ala Glu Lys Lys Lys

1 5 10 15

Lys Glu Asp Pro Gly Lys Trp Gln Arg Arg Leu Tyr Val Gln Leu Phe

20 25 30

Pro Glu Gln Arg Leu Ala Gln Leu Arg Arg His Pro Ser Leu Leu Ser

35 40 45

Glu Thr Phe Pro Leu Pro Pro Ala Met Thr Met Leu Met Lys Leu Arg

50 55 60

Ser Gly Arg Arg

65

&lt;210&gt; 19

&lt;211&gt; 662

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..662:Ceres Seq. ID 1009003

&lt;400&gt;19

catttttct tctctctct cctccctaag caaaactaaa acaagctatg gctggtatgc

60

ttccggaggt tgaagtgcga agggaggcgg gcttccacgg tgggtctct cagattgaat

120

ctcgaacac agcttctgtg gcggtgcgg cgggacaagt ctggacaagg cgaccatcgt

180

tctctctta cactaccaat catgagagcc accaggccca tgtctcttc tcggagagaa

240

gigttaggaa taatctttat ggagagaca acgatgagaa acttgacgga gcagccaag

300

tggcaagac agggcttaac aagcgtcga gaatcccacc acgtacaagt tcaggcaaaa

360

agctaaagac aaagggaata aattggagca aggaaggtt aaacctctcg ggaattacc

420

gaccgaggtg gtcgggttaa agaaagccg aggaagttg atggaatggt tcaagcgcg

480

agttagggaa caacaaagatt gtgtatatg tctagaccgg ttaagaagg gtgagacatt

540

ggtcacacta caactgtccc ataagtttca ctccatgc ttatgtcctt gctgagacac

600

taatgtttat tgcctcatatt gtagaactga tatttggaaat taaatgttat atttttgatg

660

cc

662

&lt;210&gt; 20

&lt;211&gt; 130

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..130:Ceres Seq. ID 1009004

&lt;400&gt;20

Met Ala Gly Met Leu Pro Gly Val Glu Cys Ala Arg Arg Arg Arg Phe

1 5 10 15

His Gly Gly Ala Pro Pro Ile Glu Ser Ser Asn Thr Ala Ser Val Ala

20 25 30

Ala Ala Ala Gly His Val Trp Thr Arg Arg Pro Ser Phe Ser Leu Tyr

35 40 45

Thr Thr Asn His Glu Ser His Gln Ala His Val Ser Phe Ser Glu Arg

50 55 60

Ser Val Arg Asn Lys Ser Tyr Gly Glu Asp Asn Asp Glu Lys Leu Asp

662

10

SUBSTITUTE SHEET (RULE 26)

65  
Gly Ala Ala Lys Glu Ala Lys Gln Arg Ile Asn Lys Arg Leu Arg Ile 80  
85  
Pro Pro Arg Thr Ser Ser Gly Lys Met Val Lys Thr Lys Gly Ile Asn 95  
100  
Trp Ser Lys Glu Arg Val Asn Ileu Ser Gly Thr Tyr Arg Pro Arg Trp 110  
115  
Ser Gly 120  
130

<210> 21  
<211> 127  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..127:Ceres Seq. ID 1009005

<400>21  
Met Leu Pro Gly Val Glu Cys Ala Arg Arg Arg Arg Phe His Gly Gly 1  
5 10 15  
Ala Pro Pro Ile Glu Ser Ser Asn Thr Ala Ser Val Ala Ala Ala 20  
25  
Gly His Val Trp Thr Arg Arg Pro Ser Phe Ser Leu Tyr Thr Thr Asn 30  
35  
His Glu Ser His Gln Ala His Val Ser Phe Ser Glu Arg Ser Val Arg 40  
45  
Asn Lys Ser Tyr Gly Glu Asp Asn Asp Glu Lys Ileu Asp Gly Ala Ala 50  
55  
65  
Lys Glu Ala Lys Gln Arg Leu Asn Lys Arg Leu Arg Ile Pro Pro Arg 60  
65  
Thr Ser Ser Gly Lys Met Val Lys Thr Lys Gly Ile Asn Trp Ser Lys 70  
75  
Glu Arg Val Asn Leu Ser Gly Thr Tyr Arg Pro Arg Trp Ser Gly 80  
85  
115 120 125

<210> 22  
<211> 504  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..504:Ceres Seq. ID 1009345

<400>22  
aagaacaaa ggtgtcaaga agaagatcaa acatatatt tgactaaaaa tgaatgatg 60  
gatgataac ttgttgga ttgtgtgc ttgtgtgtg gagcaatcag aggtcgcga 120  
aggtcgaag tatttaac cagggtgtg tgacgggtgt cgtgtgcta atctccagc 180  
ggatgtcat ccacaaatt ccacacaaa acotgcgtc cctgttaca ataatagtg 240  
tgggtgagt agaattaccc ggtgcagac agatgcctag gtttaacct acatgaagac 300  
gaactccct tataatgac ctctctac atgagacttg caatgatatt atatatcat 360  
atatataac ataacgtgt atagtataa tccagttgat taatatata atcttactga 420  
tttaactca atacacaga tccgtgaatt gaataaatt tgttcaatgt ttgtgcttat 504  
aataatggia tgactcttat tatt

<210> 23  
<211> 76

<212> PRT  
<213> Arabidopsis thaliana  
<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..76:Ceres Seq. ID 1009346

<400>23  
Met Asn Ala Trp Met Ile Ile Leu Leu Val Ile Cys Val Ala Val Val 1  
5 10 15  
Val Glu Gln Ser Glu Ala Arg Lys Gly Arg Lys Tyr Leu Asn Pro Gly 20  
25  
Val Leu Asp Arg Cys Arg Gly Pro Asn Pro Pro Ala Gly Cys His Pro 30  
35  
His Asn Ser His His Lys Pro Arg Val Pro Val His Asn Tyr Ser Arg 40  
45  
50  
Gly Cys Ser Arg Ile Thr Arg Cys Arg Arg Asp Ala 60  
65 70 75

<210> 24  
<211> 72  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..72:Ceres Seq. ID 1009347

<400>24  
Met Ile Ile Leu Leu Val Ile Cys Val Ala Val Val Glu Gln Ser 1  
5 10 15  
Glu Ala Arg Lys Gly Arg Lys Tyr Leu Asn Pro Gly Val Leu Asp Arg 20  
25  
Cys Arg Gly Pro Asn Pro Pro Ala Gly Cys His Pro His Asn Ser His 30  
35  
His Lys Pro Arg Val Pro Val His Asn Tyr Ser Arg Gly Cys Ser Arg 40  
45  
50  
Ile Thr Arg Cys Arg Arg Asp Ala 60  
65 70

<210> 25  
<211> 967  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..967:Ceres Seq. ID 1010140

<400>25  
attttattc ttctcaacc attctcaatc ctctctctcc aatggttcc ttatctctg 60  
ttctcttc tcaactttag ctacgcactc ctctctatgc cctacgcagc actgctctt 120  
ctctcgaa ctcctgttca ttctcgta aggtctcaat cgtcgctt tcaaggacy 180  
atttgaaga gtcgcgcg gagaagctg tggaggaat taaacctggg atgttctag 240  
ggctcggaac cggatccacc cgaacttgc cigtgtgca gatcggaac ctactctt 300  
ccggtgaact ctacgatatt tgcgtatcc caactcgaa acgaacggag gaacagcac 360  
ggctgtagg gattctctt gttgggttag atacatcc gagaatcag ctgcattg 420  
acggagcaga cgaggtagat ccgaacttg attagtcaa aggtcgtgga ggtgctctc 480



tcctgagaa aatggtgaa gctgtggctg acaagttat tgtgtggct gatgatacca 540  
 aactcttcc agactcgggt ggaagtgat tagctatgcc ggtggaagtt gtccaattct 600  
 gctggaact taattgatt agattgcaag agctcttcaa ggaatttggg tgggaatcaa 660  
 agcttagagt tgaatgctat ggaagacctt atgtgactga taacagttaac tacattattg 720  
 atttgaatt taagactcct tgaagcgtg gattcgctgc ggttaagag attggaagt 780  
 ttaaggaagt ggtggaagat ggtctgttcc tgggaatggc taacttctgc attatgctg 840  
 gaaagaatg cgttgaagtt atgacaagat gaggttaaaa attcaattct tctctcttc 900  
 ttgtatgac aaattttct ccttgtaaa attatcaatt gagttttat gtgagatcat 960  
 ttcaatc 967

<210> 26  
 <211> 289

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1...289:Ceres Seq. ID 1010141

<400>26

Phe Ile Ser Ser Pro Thr Ile Leu Asn Pro Leu Ser Pro Met Ala Ser 1  
 Leu Ser Phe Val Ser Ser Ser His Leu Thr Leu Arg Thr Pro Ser Ile 20  
 Ala Leu Arg Ser Thr Gly Ser Ser Pro Arg Thr Ser Val Ser Phe Ser 35  
 Val Lys Ala Gln Sor Val Ala Leu Ser Gln Asp Asp Leu Lys Lys Leu 50  
 Ala Ala Glu Lys Ala Val Glu Ala Ile Lys Pro Gly Met Val Leu Gly 65  
 Leu Gly Thr Gly Ser Thr Ala Ala Phe Ala Val Asp Gln Ile Gly Lys 85  
 Leu Leu Ser Ser Gly Glu Leu Tyr Asp Ile Val Gly Ile Pro Thr Ser 100  
 Lys Arg Thr Glu Glu Gln Ala Arg Ser Leu Gly Ile Pro Leu Val Gly 115  
 Leu Asp Thr His Pro Arg Ile Asp Leu Ala Ile Asp Gly Ala Asp Glu 130  
 Val Asp Pro Asn Leu Asp Leu Val Lys Gly Arg Gly Gly Ala Leu Leu 145  
 Arg Glu Lys Met Val Glu Ala Val Ala Asp Lys Phe Ile Val Val Ala 165  
 Asp Asp Thr Lys Leu Val Thr Gly Leu Gly Gly Ser Gly Leu Ala Met 180  
 Pro Val Glu Val Val Gln Phe Cys Trp Asn Phe Asn Leu Ile Arg Leu 195  
 Gln Glu Leu Phe Lys Glu Phe Gly Cys Glu Ser Lys Leu Arg Val Asp 210  
 Gly Asp Gly Lys Pro Tyr Val Thr Asp Asn Ser Asn Tyr Ile Ile Asp 225  
 Leu Tyr Phe Lys Thr Pro Leu Lys Asp Gly Phe Ala Ala Lys Glu 245  
 Ile Gly Lys Phe Gln Gly Val Val Glu His Gly Leu Phe Leu Gly Met 260  
 Ala Thr Ser Val Ile Ile Ala Gly Lys Asn Gly Val Glu Val Met Thr 275  
 Lys 285

<210> 27

13

SUBSTITUTE SHEET (RULE 26)

<211> 276  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1...276:Ceres Seq. ID 1010142  
 <400>27  
 Met Ala Ser Leu Ser Phe Val Ser Ser Ser His Leu Thr Leu Arg Thr 1  
 Pro Ser Ile Ala Leu Arg Ser Thr Gly Ser Ser Pro Arg Thr Ser Val 10  
 Ser Phe Ser Val Lys Ala Gln Ser Val Ala Leu Ser Gln Asp Asp Leu 20  
 Lys Lys Leu Ala Ala Glu Lys Ala Val Glu Ala Ile Lys Pro Gly Met 35  
 Val Leu Gly Leu Gly Thr Gly Ser Thr Ala Ala Phe Ala Val Asp Gln 50  
 Ile Gly Lys Leu Leu Ser Ser Gly Glu Leu Tyr Asp Ile Val Gly Ile 65  
 Pro Thr Ser Lys Arg Thr Glu Glu Gln Ala Arg Ser Leu Gly Ile Pro 85  
 Leu Val Gly Leu Asp Thr His Pro Arg Ile Asp Leu Ala Ile Asp Gly 100  
 Ala Asp Glu Val Asp Pro Asn Leu Asp Leu Val Lys Gly Arg Gly Gly 115  
 Ala Leu Leu Arg Glu Lys Met Val Glu Ala Val Ala Asp Lys Phe Ile 130  
 Val Val Ala Asp Asp Thr Lys Leu Val Thr Gly Leu Gly Gly Ser Gly 145  
 Leu Ala Met Pro Val Glu Val Val Gln Phe Cys Trp Asn Phe Asn Leu 165  
 Ile Arg Leu Gln Glu Leu Phe Lys Glu Phe Gly Cys Glu Ser Lys Leu 180  
 Arg Val Asp Gly Asp Gly Lys Pro Tyr Val Thr Asp Asn Ser Asn Tyr 195  
 Ile Ile Asp Leu Tyr Phe Lys Thr Pro Leu Lys Asp Gly Phe Ala Ala 210  
 Ala Lys Glu Ile Gly Lys Phe Gln Gly Val Val Glu His Gly Leu Phe 225  
 Leu Gly Met Ala Thr Ser Val Ile Ile Ala Gly Lys Asn Gly Val Glu 245  
 Val Met Thr Lys 260  
 265  
 270  
 275

<210> 28

<211> 213

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1...213:Ceres Seq. ID 1010143

<400>28

Met Val Leu Gly Leu Gly Thr Gly Ser Thr Ala Ala Phe Ala Val Asp 1  
 Gln Ile Gly Lys Leu Leu Ser Ser Gly Glu Leu Tyr Asp Ile Val Gly 5  
 10  
 15

14

SUBSTITUTE SHEET (RULE 26)

20 25 30  
 Ile Pro Thr Ser Lys Arg Thr Glu Gln Ala Arg Ser Leu Gly Ile  
 35 40 45  
 Pro Leu Val Gly Leu Asp Thr His Pro Arg Ile Asp Leu Ala Ile Asp  
 50 55 60  
 Gly Ala Asp Glu Val Asp Pro Asn Leu Asp Leu Val Lys Gly Arg Gly  
 65 70 75 80  
 Gly Ala Leu Leu Arg Glu Lys Met Val Glu Ala Val Ala Asp Lys Phe  
 85 90 95  
 Ile Val Val Ala Asp Asp Thr Lys Leu Val Thr Gly Leu Gly Gly Ser  
 100 105 110  
 Gly Leu Ala Met Pro Val Glu Val Val Gln Phe Cys Tyr Asn Phe Asn  
 115 120 125  
 Leu Ile Arg Leu Gln Glu Leu Phe Lys Glu Phe Gly Cys Glu Ser Lys  
 130 135 140  
 Leu Arg Val Asp Gly Asp Gly Lys Pro Tyr Val Thr Asp Asn Ser Asn  
 145 150 155 160  
 Tyr Ile Ile Asp Leu Tyr Phe Lys Thr Pro Leu Lys Asp Gly Phe Ala  
 165 170 175  
 Ala Ala Lys Glu Ile Gly Lys Phe Gln Gly Val Val Glu His Gly Leu  
 180 185 190  
 Phe Leu Gly Met Ala Thr Ser Val Ile Ile Ala Gly Lys Asn Gly Val  
 195 200 205  
 Glu Val Met Thr Lys  
 210

&lt;210&gt; 29

&lt;211&gt; 623

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..623:Ceres Seq. ID 1010217

&lt;400&gt;29

gaaacccct aagtaagca aattcaag caattcaaaa accctagcct ctctctctt  
 atctctcaaca acaaatctt tgaatggcg atctctaagg ctctcaattgt tgtctcatg  
 atggtgatta tctccgtct tgcattcgca cagctcgagg caccagacc aagctcact  
 tctgattcta tgcgatctc agctgtcttc gtgtcagctg tgcagttatc accgcgcgag  
 aggaagcaaa ggggtttgag gtatagagag aagaggaaga atcggaatt tgaagaagc  
 atcggtatg cgtcgctaa agcttcgct gtagtgaggc cgagaatcaa aggaagttt  
 gctaagaa cagattcgag agagaatgat ggtgagagc tggagattt ttgggattc  
 ggtgtgttc cgagttctg atattccg gttaaagaa caiggtagta gtatatcgtt  
 taatagaga ttaaggataa ttaaaactaa aatgttctt gatttaacta ggggctaagt  
 taatgtaga agattgtgt ttttgaatt aattctaca attgtactt aaattatata  
 gtttacttaa tgttgacta atg

&lt;210&gt; 30

&lt;211&gt; 118

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..118:Ceres Seq. ID 1010218

&lt;400&gt;30

Met Ala Ile Ser Lys Ala Ser Ile Val Val Leu Met Met Val Ile Ile

1 10 15  
 Ser Val Val Ala Ser Ala Gln Ser Glu Ala Pro Ala Pro Ser Pro Thr  
 20 25 30  
 Ser Gly Ser Ser Ala Ile Ser Ala Ser Phe Val Ser Ala Val Gln Leu  
 35 40 45  
 Ser Pro Ala Gln Arg Glu Ala Arg Val Leu Arg Tyr Arg Glu Lys Arg  
 50 55 60  
 Lys Asn Arg Lys Phe Glu Lys Thr Ile Arg Tyr Ala Ser Arg Lys Ala  
 65 70 75 80  
 Tyr Ala Glu Met Arg Pro Arg Ile Lys Gly Arg Phe Ala Lys Arg Thr  
 90 95  
 Asp Ser Arg Glu Asn Asp Gly Gly Asp Val Gly Val Tyr Cys Gly Phe  
 100 105 110  
 Gly Val Val Pro Ser Phe  
 115

&lt;210&gt; 31

&lt;211&gt; 107

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..107:Ceres Seq. ID 1010219

&lt;400&gt;31

Met Met Val Ile Ile Ser Val Val Ala Ser Ala Gln Ser Glu Ala Pro  
 1 5 10 15  
 Ala Pro Ser Pro Thr Ser Gly Ser Ser Ala Ile Ser Ala Ser Phe Val  
 20 25 30  
 Ser Ala Val Gln Leu Ser Pro Ala Glu Arg Glu Ala Arg Val Leu Arg  
 35 40 45  
 Tyr Arg Glu Lys Arg Lys Asn Arg Lys Phe Glu Lys Thr Ile Arg Tyr  
 50 55 60  
 Ala Ser Arg Lys Ala Tyr Ala Glu Met Arg Pro Arg Ile Lys Gly Arg  
 65 70 75 80  
 Phe Ala Lys Arg Thr Asp Ser Arg Glu Asn Asp Gly Gly Asp Val Gly  
 90 95  
 Val Tyr Cys Gly Phe Gly Val Val Pro Ser Phe  
 100 105

&lt;210&gt; 32

&lt;211&gt; 106

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..106:Ceres Seq. ID 1010220

&lt;400&gt;32

Met Val Ile Ile Ser Val Val Ala Ser Ala Gln Ser Glu Ala Pro Ala  
 1 5 10 15  
 Pro Ser Pro Thr Ser Gly Ser Ser Ala Ile Ser Ala Ser Phe Val Ser  
 20 25 30  
 Ala Val Gln Leu Ser Pro Ala Glu Arg Glu Ala Arg Val Leu Arg Tyr  
 35 40 45  
 Arg Glu Lys Arg Lys Asn Arg Lys Phe Glu Lys Thr Ile Arg Tyr Ala  
 50 55 60

Ser Arg Lys Ala Tyr Ala Glu Met Arg Pro Arg Ile Lys Gly Arg Phe  
65 75  
Ala Lys Arg Thr Asp Ser Arg Glu Asn Asp Gly Gly Asp Val Gly Val  
85 95  
Tyr Cys Gly Phe Gly Val Val Pro Ser Phe  
100 105

&lt;210&gt; 33

&lt;211&gt; 568

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..568:Ceres Seq. ID 1010302

&lt;400&gt;33

aaacaaatc attcaagac atacaataa attggtttt ttttttaatt agaaacaat  
60  
gggttgagt ggtgtcttc atgtggagt tggagtaag tctcggctg aaagtctg  
120  
ggtagccctc ggtacggca tcaatctctt ccccaagct ttcctaagc actacaaac  
180  
caccagtt ctacggcgc aggaacgc tctgtgtcc attgctca ttacttatg  
240  
agaagatct cactgttga agatctgcg ggaaggatc gaagcagtg attggaga  
300  
caaaagcatg tctacagca tcatgtggc ggaatgttg gagtactaca aaacgtcaa  
360  
aggaacatc accgttattc ctagaacgg tggcgctct ctgaaatgt ctgtgagtt  
420  
tgaagaacc gccatgaga ttagtgacc accgtcatc aaggacttg ctgtcaaga  
480  
cttcaagag atagatgagt atctcttcaa gcaaaactg gccaaactc agaacctta  
540  
aattatataa gagggttcga tegtctct  
568

&lt;210&gt; 34

&lt;211&gt; 155

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..155:Ceres Seq. ID 1010303

&lt;400&gt;34

Met Gly Leu Ser Gly Val Leu His Val Glu Val Glu Val Lys Ser Pro  
1 5 10 15  
Ala Glu Lys Phe Thr Val Ala Leu Gly Asp Gly Ile Asn Leu Phe Pro  
20 25 30  
Lys Ala Phe Pro Asn Asp Tyr Lys Thr Ile Gln Val Leu Ala Gly Asp  
35 40 45  
Gly Asn Ala Pro Gly Ser Ile Arg Leu Ile Thr Tyr Gly Glu Gly Ser  
50 55 60  
Pro Leu Val Lys Ile Ser Ala Glu Arg Ile Glu Ala Val Asp Leu Glu  
65 70 75 80  
Asn Lys Ser Met Ser Tyr Ser Ile Ile Gly Gly Glu Met Leu Glu Tyr  
85 90 95  
Tyr Lys Thr Phe Lys Gly Thr Ile Thr Val Ile Pro Lys Asn Gly Gly  
100 105 110  
Ser Leu Leu Lys Trp Ser Gly Glu Phe Glu Lys Thr Ala His Glu Ile  
115 120 125  
Asp Asp Pro His Val Ile Lys Asp Phe Ala Val Lys Asn Phe Lys Glu  
130 135 140  
Ile Asp Glu Tyr Leu Leu Lys Gln Thr Ser Ala  
145 150 155

<210> 35  
<211> 531  
<212> DNA  
<213> Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..531:Ceres Seq. ID 1010815

&lt;400&gt;35

ggcgccgcg aaccggaaga agaactctct ctgttaaac tgcgatctga atcttccgg  
60  
cgtgaaaaa tggttcgga gaagaaggc agaaagtcctc atgaaggat caacagaga  
120  
ttagctcttg tgaagaagc tgaagatgac accttgatg acaagctctg tctcaaatc  
180  
cttcgagct ccaaaagtaa gctgatattg atatctaga attgccacc gtlgagaag  
240  
tcagaattg aatactatgc gatcttgtt aaagtggag ttatcgcta caatggcaac  
300  
aatgtgatt tgggtactgc ttgtgtaaa tacttccag ttcttgctt cagcatcgt  
360  
gatcttggt attccagat catcaagaca ctctctggag atcagtgatt ctgatttga  
420  
tgatttgcg atgtgtctt ctatgctga ttatgcttg ttgttctcaa agacttaag  
480  
ttagattct ttggttccc gacttctct taatggttat aattaattac t  
531

&lt;210&gt; 36

&lt;211&gt; 112

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..112:Ceres Seq. ID 1010816

&lt;400&gt;36

Met Val Ala Glu Lys Lys Ala Lys Lys Ser His Gln Gly Ile Asn Ser  
1 5 10 15  
Arg Leu Ala Leu Val Met Lys Ser Gly Lys Tyr Thr Leu Gly Tyr Lys  
20 25 30  
Ser Val Leu Lys Ser Leu Arg Ser Ser Lys Gly Lys Leu Ile Leu Ile  
35 40 45  
Ser Ser Asn Cys Pro Leu Arg Arg Ser Glu Ile Glu Tyr Tyr Ala  
50 55 60  
Met Leu Ala Lys Val Gly Val His Arg Tyr Asn Gly Asn Asn Val Asp  
65 70 75  
Leu Gly Thr Ala Cys Gly Lys Tyr Phe Arg Val Ser Cys Leu Ser Ile  
80 85 90 95  
Val Asp Pro Gly Asp Ser Asp Ile Ile Lys Thr Leu Pro Gly Asp Gln  
100 105 110

&lt;210&gt; 37

&lt;211&gt; 91

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..91:Ceres Seq. ID 1010817

&lt;400&gt;37

Met Lys Ser Gly Lys Tyr Thr Leu Gly Tyr Lys Ser Val Leu Lys Ser  
1 5 10 15  
Leu Arg Ser Ser Lys Gly Lys Leu Ile Leu Ile Ser Ser Asn Cys Pro

Pro Leu Arg Arg Ser Glu Ile Glu Tyr Tyr Ala Met Leu Ala Lys Val  
 20 25 30  
 35  
 Gly Val His Arg Tyr Asn Gly Asn Val Asp Leu Gly Thr Ala Cys  
 50 55 60  
 Gly Lys Tyr Phe Arg Val Ser Cys Leu Ser Ile Val Asp Pro Gly Asp  
 65 70 75 80  
 Ser Asp Ile Ile Lys Thr Leu Pro Gly Asp Gln  
 85 90

&lt;210&gt; 38

&lt;211&gt; 73

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..73:Ceres Seq. ID 1010818

&lt;400&gt;38

Met Arg Cys Leu Leu Lys Leu Glu Phe Ile Ala Thr Met Ala Thr Met  
 1 5 10 15  
 Leu Ile Trp Val Leu Leu Val Val Asn Thr Ser Glu Phe Leu Ala Ser  
 20 25 30  
 Ala Ser Leu Ile Leu Val Ile Pro Thr Ser Arg His Phe Leu Glu  
 35 40 45  
 Ile Ser Asp Ser Asp Phe Asp Asp Phe Ala Met Leu Phe Ser Met Leu  
 50 55 60  
 Asp Tyr Ala Leu Cys Ser Ser Lys Thr  
 65 70

&lt;210&gt; 39

&lt;211&gt; 494

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..494:Ceres Seq. ID 1011437

&lt;400&gt;39

acaatcgaa gaaaggagaa gacgaagatg tcgttggtat ggcgtgaagc gatgtgcct  
 60  
 ctggaaatca tcgtgggat gctctgtatc atgggaatt ctacgtacta catccacaaa  
 120  
 gcttatcatg gcgttcctaa gcaatcggc cagcatgaat gggatgttc tatggaaga  
 180  
 cgcgaaga aagctcgtga gaaagtcga gctcttcct catgattcgc ttatctctt  
 240  
 ttgttccct caggggctta agtgaccac ttgtgtgac aaataaagtg cattccagaa  
 300  
 gaagaagaag ctgggggagc tagtaattc attcccaatt gatttccct ggacatatta  
 360  
 aagcttcag aaatcagacc toataaacat ttggtttatc aaattattc tatcgtgat  
 420  
 atttgtatg ctttaagt gcatgataca ctgaactact cctctgctgt gtgagtgaat  
 480  
 aaatgaatat ctgt  
 494

&lt;210&gt; 40

&lt;211&gt; 94

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..94:Ceres Seq. ID 1011438

&lt;400&gt;40

Gln Ser Glu Glu Arg Arg Arg Arg Cys Arg Trp Tyr Gly Trp Lys  
 1 5 10 15  
 Arg Cys Cys Leu Ser Glu Ser Val Gly Cys Ser Val Ser Trp Ala  
 20 25 30  
 Ile Leu Ser Thr Thr Ser Thr Lys Leu Ile Met Ala Val Leu Ser Thr  
 35 40 45  
 Ser Ala Thr Met Asn Gly Met Phe Leu Trp Lys Asp Ala Thr Arg Lys  
 50 55 60  
 Ser Ser Arg Lys Leu Gln Leu Leu Pro His Asp Ser Leu Tyr Leu Phe  
 65 70 75 80  
 Cys Val Pro Gln Gly Leu Lys Val Thr Cys Gly Asp Lys  
 85 90

&lt;210&gt; 41

&lt;211&gt; 74

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..74:Ceres Seq. ID 1011439

&lt;400&gt;41

Thr Ile Gly Arg Lys Glu Lys Thr Lys Met Ser Leu Val Trp Leu Glu  
 1 5 10 15  
 Ala Met Leu Pro Leu Gly Ile Ile Gly Gly Met Leu Cys Ile Met Gly  
 20 25 30  
 Asn Ser Gln Tyr Tyr Ile His Lys Ala Tyr His Gly Arg Pro Lys His  
 35 40 45  
 Ile Gly His Asp Glu Trp Asp Val Ser Met Glu Arg Arg Asp Lys Lys  
 50 55 60  
 Val Val Glu Lys Ala Ala Pro Ser Ser  
 65 70

&lt;210&gt; 42

&lt;211&gt; 65

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..65:Ceres Seq. ID 1011440

&lt;400&gt;42

Met Ser Leu Val Trp Leu Glu Ala Met Leu Leu Pro Leu Gly Ile Ile Gly  
 1 5 10 15  
 Gly Met Leu Cys Ile Met Gly Asn Ser Gln Tyr Tyr Ile His Lys Ala  
 20 25 30  
 Tyr His Gly Arg Pro Lys His Ile Gly His Asp Glu Trp Asp Val Ser  
 35 40 45  
 Met Glu Arg Arg Asp Lys Lys Val Val Glu Lys Ala Ala Pro Ser  
 50 55 60  
 Ser  
 65

&lt;210&gt; 43

<211> 1011  
 <212> DNA  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..1011:Ceres Seq. ID 1011616

<400>43  
 ctgtttcttt ttgttctctt agactctcc gtccattaac ctttgcgag ggtcattttt  
 ttttttgaa ttttttacc aagtttgtgt tcaaggattt gctgtctctc cagtatgagt  
 atctccggtg ctgcagtttg tccagcaga aattaaaga ggcgggtaga gtttggaaa  
 actcatgttg ttggcctaa agggaaacat caagcaacta ttgtctggtt acatgggtt  
 ggggacaatg gtcgagctg gtccagattt ttggagacc ttccctctcc aaatatcaaa  
 tggatttgc cgaactgtcc ttccaacca ataaagtatt ttgttggttt tccctccaa  
 gcttggttg atgtgtgga caactatgaa gatggacctg atgatatgga aggattggat  
 gtggtctg ccatgttgc aaactgtgt tcaatgagc ctgctgacat taaattaggt  
 gtggaggt tcaagctggg tgcggcgaca tctctatatt ctgcaacttg tttgtctc  
 ggttaaatg gaaatgcaa tccatccct atcaatttaa ggcgaatcat aggtctaaagc  
 ggtggcttc cttgtgaaa cactattgct ggcaactag aagaggaaca gatcaagaac  
 cgactgcat cgttaccat tgttcttgt catggaaaag ctgatgatgt ggtaccgttc  
 aagttgggg agaaacttc acaggtcttg cttccaagt ggttlaaga ggtgaccttc  
 aaacttaca gtgcacttg tcaccaca atcccacag agttgatatg gttgtgcga  
 tggtagcat ccagctcag cctcaaggt tgaactctc tatgatgtag tttctgatg  
 aaaaacctc aactcttag agttgatg aattgatgt ttcaggattt cacaatgttt  
 tcaattgcat attgtgaa gacacattcg tctagtatga aattctctcc c  
 1011

<210> 44  
 <211> 252  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..252:Ceres Seq. ID 1011617

<400>44  
 Met Ser Ile Ser Gly Ala Ala Val Gly Ser Gly Arg Asn Leu Arg Arg  
 1 5 10 15  
 Ala Val Glu Phe Gly Lys Thr His Val Val Arg Pro Lys Gly Lys His  
 20 25 30  
 Gln Ala Thr Ile Val Trp Leu His Gly Leu Gly Asp Asn Gly Ser Ser  
 35 40 45  
 Trp Ser Gln Leu Leu Glu Thr Leu Pro Leu Pro Asn Ile Lys Trp Ile  
 50 55 60  
 Cys Pro Thr Ala Pro Ser Gln Pro Ile Ser Leu Phe Gly Gly Phe Pro  
 65 70 75 80  
 Ser Thr Ala Trp Phe Asp Val Val Asp Ile Asn Glu Asp Gly Pro Asp  
 85 90 95  
 Asp Met Glu Gly Leu Asp Val Ala Ala His Val Ala Asn Leu Leu  
 100 105 110  
 Ser Asn Glu Pro Ala Asp Ile Lys Leu Gly Val Gly Gly Phe Ser Met  
 115 120 125  
 Gly Ala Ala Thr Ser Leu Tyr Ser Ala Thr Cys Phe Ala Leu Gly Lys  
 130 135 140  
 Tyr Gly Asn Gly Asn Pro Tyr Pro Ile Asn Leu Ser Ala Ile Ile Gly  
 145 150 155 160  
 Leu Ser Gly Trp Leu Pro Cys Ala Lys Thr Leu Ala Gly Lys Leu Glu  
 165 170 175

Glu Glu Gln Ile Lys Asn Arg Ala Ala Ser Leu Pro Ile Val Val Cys  
 180 185 190  
 His Gly Lys Ala Asp Asp Val Val Pro Phe Lys Phe Gly Glu Lys Ser  
 195 200 205  
 Ser Gln Ala Leu Leu Ser Asn Gly Phe Lys Lys Val Thr Phe Lys Pro  
 210 215 220  
 Tyr Ser Ala Leu Gly His His Thr Ile Pro Gln Glu Leu Asp Glu Leu  
 225 230 235 240  
 Cys Ala Trp Leu Thr Ser Thr Leu Ser Leu Glu Gly  
 245 250

<210> 45  
 <211> 155  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..155:Ceres Seq. ID 1011618

<400>45  
 Met Glu Gly Leu Asp Val Ala Ala His Val Ala Asn Leu Leu Ser  
 1 5 10 15  
 Asn Glu Pro Ala Asp Ile Lys Leu Gly Val Gly Gly Phe Ser Met Gly  
 20 25 30  
 Ala Ala Thr Ser Leu Tyr Ser Ala Thr Cys Phe Ala Leu Gly Tyr  
 35 40 45  
 Gly Asn Gly Asn Pro Tyr Pro Ile Asn Leu Ser Ala Ile Ile Gly Leu  
 50 55 60  
 Ser Gly Trp Leu Pro Cys Ala Lys Thr Leu Ala Gly Lys Leu Glu Glu  
 65 70 75 80  
 Glu Gln Ile Lys Asn Arg Ala Ala Ser Leu Pro Ile Val Val Cys His  
 85 90 95  
 Gly Lys Ala Asp Val Val Pro Phe Lys Phe Gly Glu Lys Ser Ser  
 100 105 110  
 Gln Ala Leu Leu Ser Asn Gly Phe Lys Lys Val Thr Phe Lys Pro Tyr  
 115 120 125  
 Ser Ala Leu Gly His His Thr Ile Pro Gln Glu Leu Asp Glu Leu Cys  
 130 135 140  
 Ala Trp Leu Thr Ser Thr Leu Ser Leu Glu Gly  
 145 150 155

<210> 46  
 <211> 643  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..643:Ceres Seq. ID 1011631

<400>46  
 aaatgtgat ggagatggga gtgaatgaga tgaatgttgg gattaaagaa gagaagaag  
 60  
 aagaagttct tatgtttaga ttgactatg aatcagtcat ttccactgg gagggcaag  
 120  
 ggatcccatg gacggccgg ttgcactcg aatagacct cgacatggt ttgttccaa  
 180  
 ccataccat ggtgaaagt ggagcagag ctcatcata caaccactc cagcgctag  
 240  
 gttaccact agggatgctt ggggatggg gaagagagc taggtttca agatccag  
 300  
 aaaaagag gacaaggtt ttctccaa agataagta cgaggtacgt aaattgaatg  
 360  
 cagaanaag gctcgcgatg aaaggaaggt tcttcaagag atcttcaatt ggtgtgtc  
 420

actaaagaac ttaattaatt atgatatatta aattactttg ctctcatctt gctttttgt 480  
 tgcatagtt ttgtgtattg ttagctttct tttctgcac toatagagaa ttgtgcagt 540  
 ttgtggagc tagctatga cataataata ttaccaaaaa aatgtgacta tctgtgaagc 600  
 actgatttat atagtcgata taactgtaac ttgtattgct ggc 643

<210> 47  
 <211> 140  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..140:Ceres Seq. ID 1011632

<400>47  
 Asn Val Met Glu Met Gly Val Asn Glu Met Ser Gly Gly Ile Lys Glu 15  
 1 5 10  
 Glu Lys Lys Glu Lys Ala Leu Met Leu Arg Leu Asp Tyr Glu Ser Val 30  
 20 25  
 Ile Ser Thr Trp Gly Gly Ile Pro Trp Thr Ala Arg Val Pro 45  
 35 40  
 Ser Glu Ile Asp Leu Asp Met Val Cys Phe Pro Thr His Thr Met Gly 60  
 50 55  
 Glu Ser Gly Ala Glu Ala His His Asn His Phe Arg Gly Leu Gly 80  
 65 70 75  
 Leu His Leu Gly Asp Ala Gly Asp Gly Gly Arg Glu Ala Arg Val Ser 95  
 85  
 Arg Tyr Arg Glu Lys Arg Arg Thr Arg Leu Phe Ser Lys Lys Ile Arg 110  
 100 105  
 Tyr Glu Val Arg Lys Leu Asn Ala Glu Lys Arg Pro Arg Met Lys Gly 125  
 115  
 Arg Phe Val Lys Arg Ser Ile Gly Val Ala His 140  
 130 135

<210> 48  
 <211> 138  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..138:Ceres Seq. ID 1011633

<400>48  
 Met Glu Met Gly Val Asn Glu Met Ser Gly Gly Ile Lys Glu Glu Lys 15  
 1 5 10  
 Lys Glu Lys Ala Leu Met Leu Arg Leu Asp Tyr Glu Ser Val Ile Ser 30  
 20 25  
 Thr Trp Gly Gly Glu Ile Pro Trp Thr Ala Arg Val Pro Ser Glu 45  
 35 40  
 Ile Asp Leu Asp Met Val Cys Phe Pro Thr His Thr Met Gly Glu Ser 60  
 50 55  
 Gly Ala Glu Ala His His Asn His Phe Arg Gly Leu Gly Leu His 75  
 65 70  
 Leu Gly Asp Ala Gly Asp Gly Arg Glu Ala Arg Val Ser Arg Tyr 95  
 85 90  
 Arg Glu Lys Arg Arg Thr Arg Leu Phe Ser Lys Lys Ile Arg Tyr Glu 110  
 100 105  
 Val Arg Lys Leu Asn Ala Glu Lys Arg Pro Arg Met Lys Gly Arg Phe 140

115 120 125  
 Val Lys Arg Ser Ile Gly Val Ala His 135  
 130

<210> 49  
 <211> 136  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..136:Ceres Seq. ID 1011634

<400>49  
 Met Gly Val Asn Glu Met Ser Gly Gly Ile Lys Glu Glu Lys Lys Glu 15  
 1 5 10  
 Lys Ala Leu Met Leu Arg Leu Asp Tyr Glu Ser Val Ile Ser Thr Trp 30  
 20 25  
 Gly Gly Glu Gly Ile Pro Trp Thr Ala Arg Val Pro Ser Glu Ile Asp 45  
 35 40  
 Leu Asp Met Val Cys Phe Pro Thr His Thr Met Gly Glu Ser Gly Ala 60  
 50 55  
 Glu Ala His His Asn His Phe Arg Gly Leu Gly Leu His Leu Gly 80  
 65 70 75  
 Asp Ala Gly Asp Gly Arg Glu Ala Arg Val Ser Arg Tyr Arg Glu 95  
 85  
 Lys Arg Arg Thr Arg Leu Phe Ser Lys Lys Ile Arg Tyr Glu Val Arg 110  
 100 105  
 Lys Leu Asn Ala Glu Lys Arg Pro Arg Met Lys Gly Arg Phe Val Lys 125  
 115  
 Arg Ser Ser Ile Gly Val Ala His 135  
 130

<210> 50  
 <211> 663  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..663:Ceres Seq. ID 1011714

<400>50  
 gtgttcgata ataattgtga caggagaagc gataccaaga gtgcgcgtcg tegtttcat 60  
 tctcaagga aactcaatct tattagtcg ccgcgttcc tcaatcgga actccacttt 120  
 cgtcttccc ggtggccacc tgaattcgg agagacttt gaagaatg cagcgagaga 180  
 agtaatggag gaacacggtc taaagattga aaagatgaag ctittgactg ttacaacaa 240  
 tgtcttcaaa gaagaccaa cgcctacaa ctacgttct ctitcgatc gtgggtgtt 300  
 ggtggatcca agtcaagaac cgaagaatat ggaaccagag aagtgatga gatggattg 360  
 gtatgattgg gagaatctac caaagcttt gtttggcca cttagaagt tgttgaag 420  
 tggttcaat ccttcaatc atgggtggg agactaatag atgtaagat taatgatga 480  
 ttgggattg aatgttcaa attgggcatt tggctagtg gtatgattc cgttaggt 540  
 gcagaggtc cagagtcaa ttctcggaat gccctctct ttactcttt ttattgtat 600  
 ttttggtttc tgtttccta attagtaata aactatacat gtgtcctatc gattttatt 660  
 ctc 663

<210> 51  
 <211> 151  
 <212> PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..151:Ceres Seq. ID 1011715

&lt;400&gt;51

Cys Ser Ile Ile Met Ser Thr Gly Glu Ala Ile Pro Arg Val Ala Val  
1 5 10 15  
Val Val Phe Ile Leu Asn Gly Asn Ser Ile Leu Leu Gly Arg Arg  
20 25 30  
Ser Ser Ile Gly Asn Ser Thr Phe Ala Leu Pro Gly Gly His Leu Glu  
35 40 45  
Phe Gly Glu Ser Phe Glu Glu Cys Ala Ala Arg Glu Val Met Glu Glu  
50 55 60  
Thr Gly Leu Lys Ile Glu Lys Met Lys Leu Leu Thr Val Thr Asn Asn  
65 70 75  
Val Phe Lys Glu Ala Pro Thr Pro Ser His Tyr Val Ser Val Ser Ile  
85 90 95  
Arg Ala Val Leu Val Asp Pro Ser Gln Glu Pro Lys Asn Met Glu Pro  
100 105 110  
Glu Lys Cys Glu Gly Trp Asp Trp Tyr Asp Trp Glu Asn Leu Pro Lys  
115 120 125  
Pro Leu Phe Trp Pro Leu Glu Lys Leu Phe Gly Ser Gly Phe Asn Pro  
130 135 140  
Phe Thr His Gly Gly Asp  
145 150

&lt;210&gt; 52

&lt;211&gt; 147

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..147:Ceres Seq. ID 1011716

&lt;400&gt;52

Met Ser Thr Gly Glu Ala Ile Pro Arg Val Ala Val Val Phe Ile  
1 5 10 15  
Leu Asn Gly Asn Ser Ile Leu Leu Gly Arg Arg Ser Ile Gly  
20 25 30  
Asn Ser Thr Phe Ala Leu Pro Gly Gly His Leu Glu Phe Gly Glu Ser  
35 40 45  
Phe Glu Glu Cys Ala Ala Arg Glu Val Met Glu Glu Thr Gly Leu Lys  
50 55 60  
Ile Glu Lys Met Lys Leu Leu Thr Val Thr Asn Asn Val Phe Lys Glu  
65 70 75 80  
Ala Pro Thr Pro Ser His Tyr Val Ser Val Ser Ile Arg Ala Val Leu  
85 90 95  
Val Asp Pro Ser Gln Glu Pro Lys Asn Met Glu Pro Glu Lys Cys Glu  
100 105 110  
Gly Trp Asp Trp Tyr Asp Trp Glu Asn Leu Pro Lys Pro Leu Phe Trp  
115 120 125  
Pro Leu Glu Lys Leu Phe Gly Ser Gly Phe Asn Pro Phe Thr His Gly  
130 135 140  
Gly Gly Asp  
145 150

&lt;210&gt; 53

&lt;211&gt; 90

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..90:Ceres Seq. ID 1011717

&lt;400&gt;53

Met Glu Glu Thr Gly Leu Lys Ile Glu Lys Met Lys Leu Leu Thr Val  
1 5 10 15  
Thr Asn Asn Val Phe Lys Glu Ala Pro Thr Pro Ser His Tyr Val Ser  
20 25 30  
Val Ser Ile Arg Ala Val Leu Val Asp Pro Ser Gln Glu Pro Lys Asn  
35 40 45  
Met Glu Pro Glu Lys Cys Glu Gly Trp Asp Trp Tyr Asp Trp Glu Asn  
50 55 60  
Leu Pro Lys Pro Leu Phe Trp Pro Leu Glu Lys Leu Phe Gly Ser Gly  
65 70 75 80  
Phe Asn Pro Phe Thr His Gly Gly Asp  
85 90

&lt;210&gt; 54

&lt;211&gt; 524

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..524:Ceres Seq. ID 1011784

&lt;400&gt;54

aaacagaac aaacacaac acaacagag atcaacgaa gaagaaaaat ggtgaaggcg  
atgtgggtt cegtttagc tctggcgag gcatctcc tttagcggg cccgtggga  
gaagggtga cgtgtcgcc tatgcagctg gcttcattcg cggggcgat gccgtatct  
tcgcgcgat cggaggcgtg ttgcacaag agcagagagc agcagcattg cctttgtggg  
240 245  
tacatgagga accctacct ccgccaatcc tttagctccc ctacgccaag gaagcttcc  
300 305  
aacagtgcga agattcttc cccaagctgt taaggaaatg ttaatcatga ttaattagtg  
360 365  
acaagtttcg cgtattatg tggtaatgc tggcttatc ttcgtcagta ctactataa  
420 425  
taataatga gtgatgata ttgatgata tcatcatcat gcttgaatgt tatgtttct  
480 485  
atctataaa abataaata tactttagea atatgtttcg tcgt  
524 529

&lt;210&gt; 55

&lt;211&gt; 110

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..110:Ceres Seq. ID 1011785

&lt;400&gt;55

Lys Thr Glu Gln Thr Gln Thr Gln Thr Glu Asn Gln Arg Arg Arg Lys  
1 5 10 15  
Met Val Lys Val Met Trp Val Ser Val Leu Ala Leu Ala Ala Ile  
20 25 30  
Leu Leu Leu Thr Val Pro Val Ala Glu Gly Val Thr Cys Ser Pro Met

35 45  
Gln Leu Ala Ser Cys Ala Ala Met Thr Ser Ser Pro Pro Ser  
50 60  
Glu Ala Cys Cys Thr Lys Leu Arg Glu Gln Pro Cys Lys Gly  
65 75  
Tyr Met Arg Asn Pro Thr Leu Arg Gln Tyr Val Ser Pro Asn Ala  
85 90 95  
Arg Lys Val Ser Asn Ser Cys Lys Ile Pro Ser Pro Ser Cys  
100 105 110

&lt;210&gt; 56

&lt;211&gt; 94

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..94:Ceres Seq. ID 1011786

&lt;400&gt;56

Met Val Lys Val Met Trp Val Ser Val Leu Ala Leu Ala Ala Ile  
1 10 15  
Leu Leu Leu Thr Val Pro Val Ala Glu Gly Val Thr Cys Ser Pro Met  
20 25 30  
Gln Leu Ala Ser Cys Ala Ala Met Thr Ser Ser Pro Pro Ser  
35 40 45  
Glu Ala Cys Cys Thr Lys Leu Arg Glu Gln Pro Cys Lys Gly  
50 55 60  
Tyr Met Arg Asn Pro Thr Leu Arg Gln Tyr Val Ser Pro Asn Ala  
65 70 75 80  
Arg Lys Val Ser Asn Ser Cys Lys Ile Pro Ser Pro Ser Cys  
85 90

&lt;210&gt; 57

&lt;211&gt; 90

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..90:Ceres Seq. ID 1011787

&lt;400&gt;57

Asn Arg Thr Asn Thr Asn Thr Asn Arg Glu Ser Thr Lys Lys Asn  
1 5 10 15  
Gly Glu Gly Asp Val Gly Phe Arg Phe Ser Ser Gly Gly Asp Ser  
20 25 30  
Pro Phe Asp Gly Pro Gly Gly Arg Arg Gly Asp Val Leu Ala Tyr Ala  
35 40 45  
Ala Gly Phe Met Cys Gly Gly Asp Asp Val Ile Phe Ala Ala Ile Gly  
50 55 60  
Gly Val Leu His Lys Ala Glu Arg Ala Ala Met Pro Leu Trp Val  
65 70 75 80  
His Glu Glu Pro Tyr Pro Pro Ile Arg  
85 90

&lt;210&gt; 58

&lt;211&gt; 1007

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..1007:Ceres Seq. ID 1011820

&lt;400&gt;58

ccaaattctt ccattgatcg ttctttgaat ctctcgatt tagcctagg gtctgatgag 60  
cagctagcta ttccatgcc actaaatgga aagaataac taatagaaga ctgtgtgagag 120  
acagatggtt cctttgact taacacgcta atgaacggt ttctctctc taactctccc 180  
gaagcactta tttttctcgc ttctgcgcgc cctttctctc attatgatcg aatttgctg 240  
aaactgggat gtaatttagc taccataaag tcgaataatc gattggtgtt ctttgacatg 300  
ctcatggita agtttcaga tggggatcaa atggaagaca atgtgagtc agttgcgaaa 360  
ctatttcggg agatacaaga aacgttcga aagtaacaga gtgaacaag tggtaacata 420  
actgttatgg tggatgacat gctctgctg gaaatgcta tctcgggag caactcgat 480  
cagctatgg acttttga tttatgcc acattaagt ctgaagcaa ttgttcattg 540  
gtcatctca atcatgaaga tatatccg agaatgaga gacctgatt ttgtctacag 600  
atggtatgcc ttgcagatgt tgtataaag gcagagcctt tagctctgg tttagcaat 660  
gatgacatg gcaattgac ttcttgaa ttcttgaa aagatgataa gcaactcagg tagaaga 720  
tcgagaaca agttgacaa ttccaattc aagatcaag aaaaattgat cgactatttc 780  
tatctggtt gcagaagctg agattagac accgttcga gcaattagcat ggtagtga 840  
tgcattcga cagtttgtt attgctgtt aggttaaca gagacactct gttcaacatt 900  
ttacttag ttctctctt ggttgaact tggtagacta gctaaagttg aatgcttgag 960  
acacagaac ttgtagacat ggaaattag taacattcta tatcttt 1007

&lt;210&gt; 59

&lt;211&gt; 266

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..266:Ceres Seq. ID 1011821

&lt;400&gt;59

Pro Asn Phe Ser Met Asp Arg Ser Leu Asn Leu Leu Asp Leu Ala Leu  
1 5 10 15  
Gly Phe Asp Glu Gln Leu Ala Ile Pro Ser Pro Leu Asn Gly Lys Val  
20 25 30  
Ile Leu Ile Glu Asp Cys Val Glu Thr Ser Gly Ser Phe Val Leu His  
35 40 45  
Gln Leu Met Lys Arg Val Leu Ser Ser Asn Ser Ser Asp Ala Leu Ile  
50 55 60  
Phe Leu Ala Phe Ala Arg Pro Phe Ser His Tyr Asp Arg Ile Leu Arg  
65 70 75 80  
Lys Leu Gly Cys Asn Leu Ala Thr His Lys Ser Asn Asn Arg Leu Val  
90 95  
Phe Phe Asp Met Leu Met Val Lys Cys Ser Asp Gly Asp Gln Met Glu  
100 105 110  
Asp Asn Val Ser Ala Val Ala Lys Leu Phe Arg Glu Ile Gln Glu Thr  
115 120 125  
Val Arg Lys Leu Gln Ser Val Thr Ser Gly Asn Ile Thr Val Met Val  
130 135 140  
Asp Asp Met Ser Leu Leu Glu Ile Ala Thr Thr Gly Ser Asn Ser Asp  
145 150 155  
His Val Leu Asp Phe Leu His Tyr Cys His Thr Leu Ser Ser Glu Ser  
165 170 175  
Asn Cys Ser Leu Val Ile Leu Asn His Glu Asp Ile Tyr Ala Ser Met  
180 185 190



Glu Arg Pro Ala Phe Ieu Leu Gln Met Val Cys Leu Ala Asp Val Val  
195 200 205  
Ile Lys Ala Glu Pro Leu Ala Ser Gly Leu Ala Asn Asp Val His Gly  
210 215 220  
Gln Ieu Thr Val Leu Asn Lys Gly Ile Ser Asn Ser Gly Arg Gly Ser  
225 230 235 240  
Ser Arg Asn Lys Leu Gln Asn Phe Gln Phe Arg Ile Lys Glu Asn Gly  
245 250 255  
Ile Asp Tyr Phe Tyr Pro Gly Cys Arg Ser  
260 265

&lt;210&gt; 60

&lt;211&gt; 262

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..262:Ceres Seq. ID 1011822

<400>60  
Met Asp Arg Ser Ieu Asn Leu Ieu Asp Leu Ala Leu Gly Phe Asp Glu  
1 5 10 15  
Gln Ieu Ala Ile Pro Ser Pro Leu Asn Gly Lys Val Ile Leu Ile Glu  
20 25 30  
Asp Cys Val Glu Thr Ser Gly Ser Phe Val Leu His Gln Ieu Met Lys  
35 40 45  
Arg Val Leu Ser Ser Asn Ser Asp Ala Leu Ile Phe Leu Ala Phe  
50 55 60  
Ala Arg Pro Phe Ser His Tyr Asp Arg Ile Leu Arg Lys Leu Gly Cys  
65 70 75 80  
Asn Ieu Ala Thr His Lys Ser Asn Asn Arg Ieu Val Phe Phe Asp Met  
85 90 95

Ieu Met Val Lys Cys Ser Asp Gly Asp Gln Met Glu Asp Asn Val Ser  
100 105 110  
Ala Val Ala Lys Leu Phe Arg Glu Ile Gln Glu Thr Val Arg Lys Leu  
115 120 125  
Gln Ser Val Thr Ser Gly Asn Ile Thr Val Met Val Asp Asp Met Ser  
130 135 140  
Leu Ieu Glu Ile Ala Thr Thr Gly Ser Asn Ser Asp His Val Leu Asp  
145 150 155  
Phe Ieu His Tyr Cys His Thr Leu Ser Ser Glu Ser Asn Cys Ser Leu  
165 170 175  
Val Ile Leu Asn His Glu Asp Ile Tyr Ala Ser Met Glu Arg Pro Ala  
180 185 190

Phe Leu Leu Gln Met Val Cys Leu Ala Asp Val Ile Lys Ala Glu  
195 200 205  
Pro Leu Ala Ser Gly Leu Ala Asn Asp Val His Gly Gln Leu Thr Val  
210 215 220  
Leu Asn Lys Gly Ile Ser Asn Ser Gly Arg Gly Ser Ser Arg Asn Lys  
225 230 235 240  
Leu Gln Asn Phe Gln Phe Arg Ile Lys Glu Asn Gly Ile Asp Tyr Phe  
245 250 255  
Tyr Pro Gly Cys Arg Ser  
260

&lt;210&gt; 61

&lt;211&gt; 216

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;210&gt; 61

&lt;211&gt; 216

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..216:Ceres Seq. ID 1011823

<400>61  
Met Lys Arg Val Leu Ser Ser Asn Ser Ser Asp Ala Leu Ile Phe Leu  
1 5 10 15  
Ala Phe Ala Arg Pro Phe Ser His Tyr Asp Arg Ile Leu Arg Lys Leu  
20 25 30  
Gly Cys Asn Leu Ala Thr His Lys Ser Asn Asn Arg Leu Val Phe Phe  
35 40 45  
Asp Met Leu Met Val Lys Cys Ser Asp Gly Asp Gln Met Glu Asp Asn  
50 55 60  
Val Ser Ala Val Ala Lys Ieu Phe Arg Glu Ile Gln Glu Thr Val Arg  
65 70 75 80  
Lys Leu Gln Ser Val Thr Ser Gly Asn Ile Thr Val Met Val Asp Asp  
85 90 95

Met Ser Ieu Ieu Glu Ile Ala Thr Thr Gly Ser Asn Ser Asp His Val  
100 105 110  
Leu Asp Phe Leu His Tyr Cys His Thr Leu Ser Ser Glu Ser Asn Cys  
115 120 125  
Ser Leu Val Ile Leu Asn His Glu Asp Ile Tyr Ala Ser Met Glu Arg  
130 135 140  
Pro Ala Phe Leu Leu Gln Met Val Cys Leu Ala Asp Val Val Ile Lys  
145 150 155 160  
Ala Glu Pro Leu Ala Ser Gly Leu Ala Asn Asp Val His Gly Gln Leu  
165 170 175  
Thr Val Leu Asn Lys Gly Ile Ser Asn Ser Gly Arg Gly Ser Ser Arg  
180 185 190  
Asn Lys Leu Gln Asn Phe Gln Phe Arg Ile Lys Glu Asn Gly Ile Asp  
195 200 205  
Tyr Phe Tyr Pro Gly Cys Arg Ser  
210 215

<210> 62  
<211> 384  
<212> DNA  
<213> Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..584:Ceres Seq. ID 1011874

<400>62  
tcaattaaca taaccttcgc cgtcgtatgaa gctttccott caattccac taccttcagc  
gtcgcacga aacaaacta cgcagtaaaa caatccacc atgaagtta tgcggagaa  
aggaagatat acgacatcag ccaccagtag acgcggagt tgcgggttg ggaacttca  
gaagactag ggaacttct tagactgcc gtgagtatga aqaatggatc cgaatctta  
actcgaaga tgaactatc tgttacctt ggaactcatg ttgatgacc aggcatttc  
catgaccatt atatagatc tggttttgat actgattcac ttgacttca aatcttaat  
ggtaaacatt ttggaaact tttttttt tttttttt gtatctatg atcgttttg  
aaattggttt ctgtgaattt aattatagg tccctcttta ttggttatg ttccaagaa  
taagaacatt tcaaggtttt tgaagatttt ctaacattga aattttaca aaacttcat  
taattggttg ttacaaatt ttgtgttaatt gatttggttg ttgt

&lt;210&gt; 63

&lt;211&gt; 135

&lt;212&gt; PRT

<210> 63  
<211> 135  
<212> PRT

&lt;213&gt; Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..135:Ceres Seq. ID 1011875

<400>63  
Ser Ile Asn Ile Thr Phe Ala Val Asp Glu Ala Phe Pro Ser Ile Pro1 5 10 15  
Thr Thr Phe Ser Val Ala Thr Lys Gln His Tyr Asp Val Lys Pro Ile20 25 30  
His His Glu Val Tyr Asp Gly Glu Arg Lys Ile Tyr Asp Ile Ser His35 40 45  
Gln Tyr Thr Pro Glu Leu Pro Val Trp Glu Ser Ser Glu Gly Leu Gly50 55 60  
Asn Phe Leu Arg Leu Ala Val Ser Met Lys Asn Gly Ser Asp Ala Asn65 70 75 80  
Ile Ser Lys Met Glu Leu Ser Val His Ser Gly Thr His Val Asp Ala85 90  
Pro Gly His Phe His Asp His Tyr Tyr Glu Ser Gly Phe Asp Thr Asp100 105 110  
Ser Leu Asp Leu Gln Ile Leu Asn Gly Lys His Phe Thr Asn Ser Phe115 120 125  
Phe Phe Phe Phe Leu Ser Ser

130 135

&lt;210&gt; 64

&lt;211&gt; 626

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..626:Ceres Seq. ID 1011981

&lt;400&gt;64

acatctaat cgaataaac octagacaaa accctctctc ctctctaat tcactttccc

60 120  
tctacaatg gcttgccga acgcagatc tgcgtcgtc tctctcgt cgcgcacgc180  
aagatacca tctctcgtt ctcacacac ggcgttctt cgcctccgtc tctctaatc240  
ctcttacct agcgtcgtc tcacttttc gtacattcc agaatctag ggcctcagg300  
ttgcacacag tgcgtcgtc acttctacag tgcgtcgtc acttctaac tgcattcca360  
ccttaacgtt aatttgccga cttctcga gtactttaa ggtacttcaa gaggcacaaag420  
agttaattt cactgcgat gcgtgaatca gtactttaa ggtacttcaa atcgaaatg480  
tggattacc taagacaaa ctgcgttaac ggtgcgtcgt ttttagaagt ttttagttca540  
caatgtggtg atttattagt tgaagagctt ttcgtatgct tgcacacatg gttttttcc600  
agagaacctt ttttgtttt ggagaaatgc atagacgat taataagtgt attggatcag626  
aactgaaca atgttgttac tgattc

&lt;210&gt; 65

&lt;211&gt; 93

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..93:Ceres Seq. ID 1011982

&lt;400&gt;65

Met Ala Trp Arg Asn Ala Gly Ser Ala Ala Arg Ser Phe Val Ser Ala

1 5 10 15  
Thr Ala Arg Ser Pro Ser Leu Arg Ser Pro Thr Thr Ala Leu Pro Arg20 25 30  
Leu Arg Pro Pro Gln Ser Ser Leu Pro Ser Arg Arg Phe Thr Phe Ser35 40 45  
Ser Pro Ser Arg Asn Leu Gly Ala Leu Gly Cys Thr Gln Ser Phe Leu50 55 60  
Pro Leu Tyr Ser Val Val Ala Thr Ser Gln Leu Thr Ser His Leu Asn65 70 75 80  
Val Asn Leu Arg Ala Phe Cys Glu Leu Ser Asn Gly Thr

85 90

&lt;210&gt; 66

&lt;211&gt; 75

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..75:Ceres Seq. ID 1011983

&lt;400&gt;66

Ile Leu Ile Glu Lys Asn Pro Ser Lys Asn Leu Ser Leu Ser Leu Asn

1 5 10 15  
Ser Leu Leu Leu Leu Gln Trp Leu Gly Ala Thr Gln Asp Leu Leu Leu20 25 30  
Val Leu Ser Ser Pro Pro Gln Asp His His Leu Ser Val Leu Pro35 40 45  
Pro Arg Arg Phe Leu Ala Ser Val Leu Leu Asn Pro Pro Tyr Leu Ala50 55 60  
Val Ala Ser Pro Ser Arg His Leu Pro Gly Ile

65 70

&lt;210&gt; 67

&lt;211&gt; 46

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..46:Ceres Seq. ID 1011984

&lt;400&gt;67

Met Cys Gly Phe Ile Ser Cys Arg Ala Phe Arg Met Leu Gly Thr Met

1 5 10 15  
Val Phe Ser Gln Arg Thr Phe Phe Cys Phe Gly Glu Met His Ser Ser20 25 30  
Ile Asn Lys Cys Ile Gly Ser Glu Leu Lys Gln Cys Cys Tyr

35 40 45

&lt;210&gt; 68

&lt;211&gt; 468

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..468:Ceres Seq. ID 1014547

<400>68  
 atagaatata accatatctc ggcacatgaag atcgtgacat tggtaactcgt cgtcttcgtc 60  
 atactttgca catcattccc ggtcccatc aaagccgaag acacgggaga tacaggaaat 120  
 gtggagatga catgtagcgc aaggcagctt cagctctgoc tggccgtgat tacgggagga 180  
 ggacacccct cgggtgcatg ttgtcaaaag cttaacagagc aacagtcagc cctatgtgtg 240  
 ttgcataaga acccgtcgct cgcacagtac attagctctc caaacgtcgc caaagtgtc 300  
 ctgtcttgc atgttgctta tccacttgt tgaaccttat ctgattttta taaataata 360  
 aacgaagga aataaattac attatatcaa acgttatgat acaattcaac cgtttgtgtt 420  
 aatgtactgg cttaactagt taaataaagt ttaatttctt ggttgctc 468

&lt;210&gt; 69

&lt;211&gt; 110

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..110:Ceres Seq. ID 1014548

&lt;400&gt;69

Ile Glu Tyr Asn His Ile Pro Asp Met Lys Ile Val Thr Leu Val Leu 15  
 1 5 10  
 Val Val Phe Val Ile Leu Ser Thr Ser Phe Pro Ala Ala Ile Lys Ala 30  
 20 25  
 Glu Asp Thr Gly Asp Thr Gly Asn Val Gly Val Thr Cys Asp Ala Arg 45  
 35 40  
 Gln Leu Gln Pro Cys Leu Ala Ala Ile Thr Gly Gly Gly Gln Pro Ser 60  
 50 55  
 Gly Ala Cys Cys Ala Lys Leu Thr Glu Gln Gln Ser Cys Leu Cys Gly 75  
 65 70  
 Phe Ala Lys Asn Pro Ala Phe Ala Gln Tyr Ile Ser Ser Pro Asn Ala 90  
 85 95  
 Arg Lys Val Leu Leu Ala Cys Asn Val Ala Tyr Pro Thr Cys 110  
 100

&lt;210&gt; 70

&lt;211&gt; 102

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..102:Ceres Seq. ID 1014549

<400>70  
 Met Lys Ile Val Thr Leu Val Leu Val Phe Val Ile Leu Ser Thr 15  
 1 5 10  
 Ser Phe Pro Ala Ala Ile Lys Ala Glu Asp Thr Gly Asp Thr Gly Asn 30  
 20 25  
 Val Gly Val Thr Cys Asp Ala Arg Gln Leu Gln Pro Cys Leu Ala Ala 45  
 35 40  
 Ile Thr Gly Gly Gln Pro Ser Gly Ala Cys Cys Ala Lys Leu Thr 60  
 50 55  
 Glu Gln Gln Ser Cys Leu Cys Gly Phe Ala Lys Asn Pro Ala Phe Ala 75  
 65 70  
 Gln Tyr Ile Ser Ser Pro Asn Ala Arg Lys Val Leu Leu Ala Cys Asn 90  
 85 95

33

SUBSTITUTE SHEET (RULE 26)

Val Ala Tyr Pro Thr Cys 100  
 100

&lt;210&gt; 71

&lt;211&gt; 1303

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..1303:Ceres Seq. ID 1014995

&lt;400&gt;71

aaaactcaag ttittgtcaa ttacctactt acaaaaagc aactgactc gaaaaaacac 60  
 atccatgct cctctggc aacctctcc tgtgagtgt toggatgaaa aatatggaaa 120  
 tgtcaagtgg gaagatgtgg cattcaagt tgtctgacg gattatagt atgttgcaa 180  
 gtgcaatcat ggagagagtt ttcaagaggg gaagattctt cctttgtcg atttgaact 240  
 taaccttgc gctgtgttc ttcaagtgg ccagggttta tatgaagAAC tgaagtcta 300  
 caggacagaa gatgtcgga ttgtctatt ccgaccagac caaaacggtc tccgcttca 360  
 agccggagct gacagactct atatgctta tccctggtc gatcaattcg tctcgccat 420  
 caaaaagtt gctctggca acaagaatg gattctcct cggggaaaag gaacattgta 480  
 tattagcct atcttgttg ggaagtgtcc gattcttgt tcaattccca tctctgagac 540  
 tgatatttg gctttgtg tctctgttg acgttatcat aagataaact cgggttgaa 600  
 cactttaca gaagatcaat ttctcgagc ttctctagt ggaactgtg gtgtgaagag 660  
 tatgaanaac tatgtccg ttgtgtgac atggcagag cggaaaaaac aaggtttc 720  
 catcaaaac tatgtccg ttgtgtgac atggcagag cggaaaaaac aaggtttc 780  
 tgatatttg gctttgtg tctctgttg caaaaacatt gaagaacttt tgcagctaa 840  
 tttttatg ctcaaggga atgtgtatc gacaccaaa attgcagaa ctattttgc 900  
 cgggtcact ggaactggc taatgaatt gtgtgtgat ttcgctacc aggttcgga 960  
 aegacgatt cctctagtg actttctga tggcgagaa gctttctga cggcactgc 1020  
 ttcattgtg acttgattg catcctaac ctttaagac aaaaagaccg gattcaaac 1080  
 aggaagaaa catgtgtgc gaagctaac gaagcgttaa gtgatatcca gacgggtcg 1140  
 gtcaagata ccaaggatg gacgtggag attgacgcc aggttgaaa gtgaactg 1200  
 taactgtg ccaaatatg tgtgtgtg atcaaaact tatgaagaa acatctgaag 1260  
 atgtctga tctttgtg tgtgtgtc atgcatgac tatcgttgc atgtattta 1303  
 taaatgttcg tctgaagtt attaatnec tatggctttt tgc

&lt;210&gt; 72

&lt;211&gt; 352

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..352:Ceres Seq. ID 1014996

&lt;400&gt;72

Lys Leu Lys Phe Leu Ala Ile Thr Tyr Leu Gln Gln Ser Asn Ser Thr 15  
 1 5 10  
 Arg Thr Asn Thr Ser Met Ala Pro Ser Ala Gln Pro Leu Pro Val Ser 30  
 20 25  
 Val Ser Asp Glu Lys Tyr Ala Asn Val Lys Trp Glu Leu Ala Phe 45  
 35 40  
 Lys Phe Val Arg Thr Asp Tyr Met Tyr Val Ala Lys Cys Asn His Gly 60  
 50 55  
 Glu Ser Phe Gln Glu Gly Lys Ile Leu Pro Phe Ala Asp Leu Gln Leu 75  
 65 70  
 Asn Pro Cys Ala Ala Val Leu Gln Tyr Gly Gln Gly Leu Tyr Glu Gly 90  
 85 95  
 Leu Lys Ala Tyr Arg Thr Glu Asp Gly Arg Ile Leu Leu Phe Arg Pro

34

SUBSTITUTE SHEET (RULE 26)

Asp Gln Asn Gly Leu Arg Leu Gln Ala Gly Ala Asp Arg Leu Tyr Met 110  
 115  
 Pro Tyr Pro Ser Val Asp Gln Phe Val Ser Ala Ile Lys Gln Val Ala 125  
 130  
 Leu Ala Asn Lys Lys Trp Ile Pro Pro Gly Lys Gly Thr Leu Tyr 140  
 145  
 Ile Arg Pro Ile Leu Phe Gly Ser Gly Pro Ile Leu Gly Ser Phe Pro 155  
 165  
 Ile Pro Glu Thr Thr Phe Thr Ala Phe Ala Cys Pro Val Gly Arg Tyr 175  
 180  
 His Lys Asp Asn Ser Gly Leu Asn Leu Lys Ile Glu Asp Gln Phe Arg 190  
 195  
 Arg Ala Phe Pro Ser Gly Thr Gly Gly Val Lys Ser Ile Thr Asn Tyr 205  
 210  
 Cys Pro Val Trp Ile Pro Leu Ala Glu Ala Lys Lys Gln Gly Phe Ser 220  
 225  
 Asp Ile Leu Phe Leu Asp Ala Ala Thr Gly Lys Asn Ile Glu Glu Leu 235  
 245  
 Phe Ala Ala Asn Val Phe Met Leu Lys Gly Asn Val Val Ser Thr Pro 255  
 260  
 Thr Ile Ala Gly Thr Ile Leu Pro Gly Val Thr Arg Asn Cys Val Met 270  
 275  
 Glu Leu Cys Arg Asp Phe Gly Tyr Gln Val Glu Glu Arg Thr Ile Pro 285  
 290  
 Leu Val Asp Phe Leu Asp Ala Asp Glu Ala Phe Cys Thr Gly Thr Ala 300  
 305  
 Ser Ile Val Thr Ser Ile Ala Ser Val Thr Phe Lys Asp Lys Lys Thr 315  
 320  
 Gly Phe Lys Thr Gly Lys Lys His Trp Leu Arg Ser Tyr Thr Arg Arg 335  
 340  
 345

&lt;210&gt; 73

&lt;211&gt; 331

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..331:Ceres Seq. ID 1014997

&lt;400&gt;73

Met Ala Pro Ser Ala Gln Pro Leu Pro Val Ser Val Ser Asp Glu Lys 1  
 5  
 Tyr Ala Asn Val Lys Trp Glu Glu Leu Ala Phe Lys Phe Val Arg Thr 10  
 20  
 Asp Tyr Met Tyr Val Ala Lys Cys Asn His Gly Glu Ser Phe Gln Glu 25  
 35  
 Gly Lys Ile Leu Pro Phe Ala Asp Leu Gln Leu Asn Pro Cys Ala Ala 40  
 50  
 Val Leu Gln Tyr Gly Gln Gly Leu Tyr Glu Gly Leu Lys Ala Tyr Arg 55  
 65  
 Thr Glu Asp Gly Arg Ile Leu Leu Phe Arg Pro Asp Gln Asn Gly Leu 70  
 85  
 Arg Leu Gln Ala Gly Ala Asp Arg Leu Tyr Met Pro Tyr Pro Ser Val 90  
 100  
 Asp Gln Phe Val Ser Ala Ile Lys Gln Val Ala Leu Ala Asn Lys Lys 105  
 115  
 Trp Ile Pro Pro Pro Gly Lys Gly Thr Leu Tyr Ile Arg Pro Ile Leu 120  
 130  
 135  
 140

Phe Gly Ser Gly Pro Ile Leu Gly Ser Phe Pro Ile Pro Glu Thr Thr 150  
 155  
 Phe Thr Ala Phe Ala Cys Pro Val Gly Arg Tyr His Lys Asp Asn Ser 160  
 165  
 Gly Leu Asn Leu Lys Ile Glu Asp Gln Phe Arg Arg Ala Phe Pro Ser 170  
 180  
 Gly Thr Gly Gly Val Lys Ser Ile Thr Asn Tyr Cys Pro Val Trp Ile 185  
 195  
 Pro Leu Ala Glu Ala Lys Lys Gln Gly Phe Ser Asp Ile Leu Phe Leu 200  
 210  
 Asp Ala Ala Thr Gly Lys Asn Ile Glu Glu Leu Phe Ala Ala Asn Val 205  
 225  
 Phe Met Leu Lys Gly Asn Val Val Ser Thr Pro Thr Ile Ala Gly Thr 235  
 245  
 Ile Leu Pro Gly Val Thr Arg Asn Cys Val Met Glu Leu Cys Arg Asp 250  
 260  
 Phe Gly Tyr Gln Val Glu Glu Arg Thr Ile Pro Leu Val Asp Phe Leu 265  
 275  
 Asp Ala Asp Glu Ala Phe Cys Thr Gly Thr Ala Ser Ile Val Thr Ser 280  
 290  
 Ile Ala Ser Val Thr Phe Lys Asp Lys Lys Thr Gly Phe Lys Thr Gly 300  
 305  
 Lys Lys His Trp Leu Arg Ser Tyr Thr Arg Arg 315  
 320  
 325  
 330

&lt;210&gt; 74

&lt;211&gt; 297

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..297:Ceres Seq. ID 1014998

&lt;400&gt;74

Met Tyr Val Ala Lys Cys Asn His Gly Glu Ser Phe Gln Glu Gly Lys 5  
 10  
 Ile Leu Pro Phe Ala Asp Leu Gln Leu Asn Pro Cys Ala Ala Val Leu 20  
 25  
 Gln Tyr Gly Gln Gly Leu Tyr Glu Gly Leu Lys Ala Tyr Arg Thr Glu 30  
 35  
 Asp Gly Arg Ile Leu Leu Phe Arg Pro Asp Gln Asn Gly Leu Arg Leu 40  
 50  
 Gln Ala Gly Ala Asp Arg Leu Tyr Met Pro Tyr Pro Ser Val Asp Gln 55  
 65  
 Phe Val Ser Ala Ile Lys Gln Val Ala Leu Ala Asn Lys Lys Trp Ile 70  
 85  
 Pro Pro Pro Gly Lys Gly Thr Leu Tyr Ile Arg Pro Ile Leu Phe Gly 90  
 100  
 Ser Gly Pro Ile Leu Gly Ser Phe Pro Ile Pro Glu Thr Thr Phe Thr 105  
 115  
 Ala Phe Ala Cys Pro Val Gly Arg Tyr His Lys Asp Asn Ser Gly Leu 120  
 130  
 Asn Leu Lys Ile Glu Asp Gln Phe Arg Arg Ala Phe Pro Ser Gly Thr 135  
 145  
 Gly Gly Val Lys Ser Ile Thr Asn Tyr Cys Pro Val Trp Ile Pro Leu 150  
 165  
 Ala Glu Ala Lys Lys Gln Gly Phe Ser Asp Ile Leu Phe Leu Asp Ala 170  
 180  
 Ala Thr Gly Lys Asn Ile Glu Glu Leu Phe Ala Ala Asn Val Phe Met 185  
 190  
 195

195  
 Leu Lys Gly Asn Val Val Sar Thr Pro Thr Ile Ala Gly Thr Ile Leu 205  
 210  
 Pro Gly Val Thr Arg Asn Cys Val Met Glu Leu Cys Arg Asp Phe Gly  
 225  
 Tyr Glu Val Glu Glu Arg Thr Ile Pro Leu Val Asp Phe Leu Asp Ala  
 245  
 Asp Glu Ala Phe Cys Thr Gly Thr Ala Ser Ile Val Thr Ser Ile Ala  
 260  
 Ser Val Thr Phe Lys Asp Lys Lys Thr Gly Phe Lys Thr Gly Lys Lys  
 275  
 His Trp Leu Arg Ser Tyr Thr Arg Arg 285  
 290

<210> 75  
 <211> 789  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..789:Ceres Seq. ID 1015323

<400>75  
 gggttaggtt tgccttcana atcgctcagaa gaagaagaag aagtcanaaga cgaatacaat 60  
 gaagtcata tgccttcacc gatttgcat gcgcgcgta gcttccctt caatacaac 120  
 catcaaac gcgcctctt gcaaaacctt ctaatactcc tcaattccc tctcaataac 180  
 ccagaanaa ctaactccg ataacgatt caactcaca cgaatcgca cgtttttac 240  
 aatccgtgt ctaatactc accacgtgtt tctttcttc gtaactgca gtaacagaa 300  
 gcaacagatc caatcggtat gttccctccc ttgatctct gataacata attcgtttg 360  
 agatgaacc agagatatt ttacgctcgt cgtcgcgct ttgctcggt ttggttggtg 420  
 tgccttaacc gctgtacta tgcattctgt ttggctctt gttggaatc gtaacagcta 480  
 tgattttag gaagaagaag agatattga gaatgatgaa tctgagctg ctgacttgaa 540  
 gaattgggt tatgttaaga ttcaagctc agctctgct cctggaagg aagctgcttg 600  
 atttgaggt gtaanttca ttgcctcat attatcata tttaactgt tgcctacta 660  
 tgggttcg tacaagtaaa ttgcctctg ttatgatgt atctaaactt actatactt 720  
 aagatcaggt gtttttaca ttgtgagttg atgtttctta tatatatata taagtgtgat 780  
 tatgatgtc 789

<210> 76  
 <211> 180  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..180:Ceres Seq. ID 1015324

<400>76  
 Met Thr Ser Ser Ser Tyr Leu Arg Phe Ala Ile Ala Val Val Ala Phe 15  
 1  
 Leu Ser Ile Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr 30  
 20  
 Ile Ser Ser Tyr Ser Leu Ser Ile Thr Thr Thr Thr Thr Thr Thr Thr 45  
 35  
 Ser Asp Phe Thr Ser Thr Arg Phe Val Thr Val Thr Thr Thr Thr Thr 60  
 50  
 Leu Asn Pro His His Val Val Val Val Val Val Val Val Val Val Val 75  
 65

Lys Pro Glu Ile Gln Ser Asp Arg Ser Leu Pro Leu Ile Ser Asp Asn 95  
 85  
 Ile Asn Ser Phe Arg Asp Arg Thr Arg Asp Ile Leu Ser Val Val Val 110  
 100  
 Ala Leu Leu Phe Gly Val Gly Cys Gly Ala Leu Thr Ala Thr Met 125  
 115  
 Tyr Leu Val Trp Ala Leu Val Val Asn Arg Gln Ser Tyr Asp Phe Glu 140  
 135  
 Glu Glu Glu Asp Asp Tyr Glu Asn Asp Glu Ser Asp Ala Ala Ser Leu 155  
 145  
 Lys Lys Leu Gly Tyr Val Lys Ile Pro Ala Pro Ala Pro Ala Pro Val 175  
 165  
 Lys Glu Ala Ala 180

<210> 77  
 <211> 495  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..495:Ceres Seq. ID 1016486

<400>77  
 attctccat tagcattcta cagtcgtcac actatacact tcttcgaaa tctctctcta 60  
 tccctcena aaatgaagat cctgtcgct ttatctcog cgtctctct cttgtctatg 120  
 gttactatg ccacaggagt gggtccagtc accgtggagg cagcacgtg tgagtcaaa 180  
 accataggt tcaagggtcc atgtgtgagc acacacaact gtgcaacgt gtgccacaac 240  
 gaagcttcg gggaggtaa atgcctgga ttccgcgtc gttgctactg cacaagacac 300  
 ttctatcca tccattctta tgactcaat tctgcacata tctgactgt ttctacttt 360  
 tctatctaa atcttcgta cgttaccata tctgacgta catgagttt ttctcaata 420  
 agcttggtt ttgtgtttt ccggttttaa tgaatgta aatcaattaa tggcttttaa 480  
 tatattgat tatgg 495

<210> 78  
 <211> 101  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..101:Ceres Seq. ID 1016487

<400>78  
 Thr Ser Ser Leu Ala Phe Tyr Ser Leu Leu His Thr Tyr His Phe Phe Arg 15  
 1  
 Asn Leu Ser Leu Ser Leu Ser Lys Met Lys Leu Ser Val Arg Phe Ile 30  
 20  
 Ser Ala Ala Leu Leu Leu Phe Met Val Phe Ile Ala Thr Gly Met Gly 45  
 35  
 Pro Val Thr Val Glu Ala Arg Thr Cys Glu Ser Lys Ser His Arg Phe 60  
 50  
 Lys Gly Pro Cys Val Ser Thr His Asn Cys Ala Asn Val Cys His Asn 75  
 65  
 Glu Gly Phe Gly Gly Lys Cys Arg Gly Phe Arg Arg Arg Cys Tyr 85  
 80  
 Cys Thr Arg His Cys 100

<210> 79  
 <211> 77  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..77:Ceres Seq. ID 1016488

<400>79  
 Met Lys Leu Ser Val Arg Phe Ile Ser Ala Ala Leu Leu Phe Met  
 1 5 10 15  
 Val Phe Ile Ala Thr Gly Met Gly Pro Val Thr Val Glu Ala Arg Thr  
 20 25 30  
 Cys Glu Ser Lys Ser His Arg Phe Lys Gly Pro Cys Val Ser Thr His  
 35 40 45  
 Asn Cys Ala Asn Val Cys His Asn Glu Gly Phe Gly Gly Lys Cys  
 50 55 60  
 Arg Gly Phe Arg Arg Cys Tyr Cys Thr Arg His Cys  
 65 70 75

<210> 80  
 <211> 62  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..62:Ceres Seq. ID 1016489

<400>80  
 Met Val Phe Ile Ala Thr Gly Met Gly Pro Val Thr Val Glu Ala Arg  
 1 5 10 15  
 Thr Cys Glu Ser Lys Ser His Arg Phe Lys Gly Pro Cys Val Ser Thr  
 20 25 30  
 His Asn Cys Ala Asn Val Cys His Asn Glu Gly Phe Gly Gly Lys  
 35 40 45  
 Cys Arg Gly Phe Arg Arg Cys Tyr Cys Thr Arg His Cys  
 50 55 60

<210> 81  
 <211> 325  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..325:Ceres Seq. ID 1018341

<400>81  
 atcatcaac acaaaacaat caatcaaaaa cataaaattc aaagtgaag caaaaccaga  
 60  
 acgagagaa atggacaaca agcaaacgc gagctacca ggcgtcaag ccactggcca  
 120  
 gactaaggag aagccggty gaatgatga caaggccaag gatgtctg cttagctca  
 180  
 agactcttg caacagactg gacaacaat gaaggagaag ggcgaaggag ctgtgtgct  
 240  
 cgtcaaggac aagaccgga tgaacaaaag caactaagca agcatcttg acttcttc  
 300  
 taattaattc cctctctcga ctgt  
 325

<210> 82  
 <211> 68  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..68:Ceres Seq. ID 1018342

<400>82  
 Met Asp Asn Lys Gln Asn Ala Ser Tyr Gln Ala Gly Gln Ala Thr Gly  
 1 5 10 15  
 Gln Thr Lys Glu Lys Ala Gly Gly Met Met Asp Lys Ala Lys Asp Ala  
 20 25 30  
 Ala Ala Ser Ala Gln Asp Ser Leu Gln Gln Thr Gly Gln Met Lys  
 35 40 45  
 Glu Lys Ala Gln Gly Ala Ala Asp Val Val Lys Asp Lys Thr Gly Met  
 50 55 60  
 Asn Lys Ser His  
 65

<210> 83  
 <211> 44  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..44:Ceres Seq. ID 1018343

<400>83  
 Met Met Asp Lys Ala Lys Asp Ala Ala Ala Ser Ala Gln Asp Ser Leu  
 1 5 10 15  
 Gln Gln Thr Gly Gln Gln Met Lys Glu Lys Ala Gln Gly Ala Ala Asp  
 20 25 30  
 Val Val Lys Asp Lys Thr Gly Met Asn Lys Ser His  
 35 40

<210> 84  
 <211> 43  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..43:Ceres Seq. ID 1018344

<400>84  
 Met Asp Lys Ala Lys Asp Ala Ala Ala Ser Ala Gln Asp Ser Leu Gln  
 1 5 10 15  
 Gln Thr Gly Gln Gln Met Lys Glu Lys Ala Gln Gly Ala Ala Asp Val  
 20 25 30  
 Val Lys Asp Lys Thr Gly Met Asn Lys Ser His  
 35 40

<210> 85  
 <211> 540  
 <212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..540:Ceres Seq. ID 1018382

<400>85

taaagcaat aaaaacaaa atggcggtt actcgaaga tcatacatg cctcattgc  
cttacgatc acatgacat ccaacagag tcaaccatc agtgtggac agtgcagac  
acattggac agtgcgtgac ctaactgac aacagtgtc cattgccatc acaatgtgc  
gtggagta agcattgta ccaattgct cagaccac cggaccgtaa acaagtatg  
gagtccta actagcggg taagaatc agggccca acaccgact tggcgcca  
cttctacca ctgtgtgtg ttaattccc taccocatca gttttagcac caattgcgac  
agatatga ctgcgtgtg aaagagcta gtgacagat gtacgactaa tcaacttgc  
cagcttttaa ccaattaaa taagaatt atgtttatat ttccatttt atgatttat  
ctcttatct atgnaacc acgatttcat atgtaataa tgacaagga tctttctc

<210> 86

<211> 119

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..119:Ceres Seq. ID 1018383

<400>86

Met Ala Phe Thr Pro Lys Ile Ile Thr Cys Leu Ile Val Leu Thr Ile  
1 5 10 15  
Tyr Met Thr Ser Pro Thr Glu Ser Thr Ile Gln Cys Gly Thr Val Thr  
20 25 30  
Ser Thr Leu Ala Gln Cys Val Thr Tyr Leu Thr Asn Ser Gly Pro Leu  
35 40 45  
Pro Ser Gln Cys Cys Val Gly Val Lys Ser Leu Tyr Gln Leu Ala Gln  
50 55 60  
Thr Thr Pro Asp Arg Lys Gln Val Cys Glu Cys Leu Lys Leu Ala Gly  
65 70 75 80  
Lys Glu Ile Lys Gly Leu Asn Thr Asp Leu Val Ala Ala Leu Pro Thr  
85 90 95  
Thr Cys Gly Val Ser Ile Pro Tyr Pro Ile Ser Phe Ser Thr Asn Cys  
100 105 110  
Asp Ser Ile Ser Thr Ala Val  
115

<210> 87

<211> 102

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..102:Ceres Seq. ID 1018384

<400>87

Met Thr Ser Pro Thr Glu Ser Thr Ile Gln Cys Gly Thr Val Thr Ser  
1 5 10 15  
Thr Leu Ala Gln Cys Val Thr Tyr Leu Thr Asn Ser Gly Pro Leu Pro  
20 25 30

Ser Gln Cys Cys Val Gly Val Lys Ser Leu Tyr Gln Leu Ala Gln Thr  
35 40 45  
Thr Pro Asp Arg Lys Gln Val Cys Glu Cys Leu Lys Leu Ala Gly Lys  
50 55 60  
Glu Ile Lys Gly Leu Asn Thr Asp Leu Val Ala Ala Leu Pro Thr Thr  
65 70 75 80  
Cys Gly Val Ser Ile Pro Tyr Pro Ile Ser Phe Ser Thr Asn Cys Asp  
85 90 95  
Ser Ile Ser Thr Ala Val  
100

<210> 88

<211> 75

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..75:Ceres Seq. ID 1018385

<400>88

Lys Ala Ile Lys Asn Lys Asn Gly Val Tyr Ser Glu Asp His His Met  
1 5 10 15  
Pro His Cys Pro Tyr Asp Leu His Asp Ile Pro Asn Arg Val Asn His  
20 25 30  
Pro Val Trp Asp Ser Asp Glu His Thr Gly Thr Val Arg Asp Leu Leu  
35 40 45  
Asp Gln Gln Trp Ser Ile Ala Ile Thr Met Leu Arg Gly Ser Gln Val  
50 55 60  
Ile Val Pro Ile Gly Ser Asp His Thr Gly Pro  
65 70 75

<210> 89

<211> 844

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..844:Ceres Seq. ID 1020666

<400>89

ctaaatattt aaagctttcc ttaccagca ttgtattgcc tccctggaac accgaagctc  
tggaataacc tatttcacc ttctaaggtt ctctcgctt ctattgatc tgtgaagatg  
tattcgggtc agacaagat ccacaaggtt aagggtgttg caccacaga gtttgaagag  
caagtactc aggtttgtt tgacttggag aacacaacc aggaattgaa aagcagttg  
aaagatctct acattacca agctgttcag atggatatct ctggcaaccg caagctgtt  
gtgactacg ttccattcga gctgagaaa gcttccgca agattcatct tegtctgtc  
agagcgttg aagaagatt cagtggaaaa gatgtatct ttgtgtac cagaagatc  
atggccccc caaagaagg ctacgtgtt cagagaccac gcaacgaac tcttactct  
gtccatgaag ccagtctcga ggaatcgtt taccctgtg agattgtagg aaagagaacc  
agataccgtc ttgatggcac caagatcatg aagttatttt tggatagtaa gctaaagac  
gacactagt acaagctgga gacaatgtg ggtgtaca gaaacttac agggaaagat  
gagtttctg agtaccattt gattctgtt ggaagtaaa ctctttgcaa agattttgtt  
ttgtcgata agagctttt gattctgtt ggaagtaaa ctctttgcaa agattttgtt  
attacagcat ttgcgaatct ctcttcaca gtccgattta tcgatacaat tctttttcc  
tacc

<210> 90

<211> 191  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..191:Ceres Seq. ID 1020667

<400>90  
 Met Tyr Ser Gly Gln Asn Lys Ile His Lys Asp Lys Gly Val Ala Pro  
 1 5 10 15  
 Thr Glu Phe Glu Gln Gln Val Thr Gln Ala Leu Phe Asp Leu Glu Asn  
 20 25 30  
 Thr Asn Gln Glu Leu Lys Ser Glu Leu Lys Asp Leu Tyr Ile Asn Gln  
 35 40 45  
 Ala Val Gln Met Asp Ile Ser Gly Asn Arg Lys Ala Val Val Ile Tyr  
 50 55 60  
 Val Pro Phe Arg Leu Arg Lys Ala Phe Arg Lys Ile His Leu Arg Leu  
 65 70 75 80  
 Val Arg Glu Leu Glu Lys Lys Phe Ser Gly Lys Asp Val Ile Phe Val  
 85 90 95  
 Ala Thr Arg Arg Ile Met Arg Pro Pro Lys Lys Gly Ser Ala Val Gln  
 100 105 110  
 Arg Pro Arg Asn Arg Thr Leu Thr Ser Val His Glu Ala Met Leu Glu  
 115 120 125  
 Asp Val Ala Tyr Pro Ala Glu Ile Val Gly Lys Arg Thr Arg Tyr Arg  
 130 135 140  
 Leu Asp Gly Thr Lys Ile Met Lys Val Phe Leu Asp Ser Lys Leu Lys  
 145 150 155 160  
 Asn Asp Thr Glu Tyr Lys Leu Glu Thr Met Val Gly Val Tyr Arg Lys  
 165 170 175  
 Leu Thr Gly Lys Asp Val Val Phe Glu Tyr Pro Val Ile Glu Ala  
 180 185 190

<210> 91  
 <211> 140  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..140:Ceres Seq. ID 1020669

<400>91  
 Met Asp Ile Ser Gly Asn Arg Lys Ala Val Val Ile Tyr Val Pro Phe  
 1 5 10 15  
 Arg Leu Arg Lys Ala Phe Arg Lys Ile His Leu Arg Leu Val Arg Glu  
 20 25 30  
 Leu Glu Lys Lys Phe Ser Gly Lys Asp Val Ile Phe Val Ala Thr Arg  
 35 40 45  
 Arg Ile Met Arg Pro Pro Lys Lys Gly Ser Ala Val Gln Arg Pro Arg  
 50 55 60  
 Asn Arg Thr Leu Thr Ser Val His Glu Ala Met Leu Glu Asp Val Ala  
 65 70 75 80  
 Tyr Pro Ala Glu Ile Val Gly Lys Arg Thr Arg Tyr Arg Leu Asp Gly  
 85 90 95  
 Thr Lys Ile Met Lys Val Phe Leu Asp Ser Lys Leu Lys Asn Asp Thr  
 100 105 110  
 Glu Tyr Lys Leu Glu Thr Met Val Gly Val Tyr Arg Lys Leu Thr Gly

SUBSTITUTE SHEET (RULE 26)

115 120 125  
 Lys Asp Val Val Phe Glu Tyr Pro Val Ile Glu Ala  
 130 135 140  
 <210> 92  
 <211> 789  
 <212> DNA  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..789:Ceres Seq. ID 1020784

<400>92  
 aggttaggt ttgtttcaa aatgctcaga agaagaaga gaagtaaaag acgataacaa  
 60  
 tgaagtcac atgctacatc cgatttcga tgcctcgtt agctttcttc tcaatcaca  
 120  
 ccatccacc cgacgtctt tgaaaacct tctaatctc ctcatattcc ctcaataaa  
 180  
 cccagaaaa cctaaacctc gaaccgatt tcaactccac acgatttcac acggtttca  
 240  
 caatcgtcg tctaaacctc caccacgtg tctctttctt cgtcaaatgt cgtcacaga  
 300  
 agccacgat ccaatcggtt cgttcctcc ctttgatctc tgataacatc aattcgttta  
 360  
 gagatcgac cagagataatt cttagcgtcg tctgcgctt ttgttcggt gttggttgg  
 420  
 gtctttaac cgtctact atgtatctg ttgggtctt tgttgtgaat cgtcaagct  
 480  
 atgatttga ggaagaaga gatgattatg agaagatga atcgtgctt gctagctga  
 540  
 agaaattgg ttatgttaag attccatctc cagctctcg tccgtgaag gaagctgctt  
 600  
 gatttaggt tgtgaattc agttcttga tattatcag attacgttg ttgcttact  
 660  
 atgtttgac gtacaagtaa atctgtctct gttatgatt tactaaact tactatact  
 720  
 taagtatacg tgttttacg attgtgatt gatgttctt atatatat ataatgtga  
 780  
 ttatgatgc 789

<210> 93  
 <211> 180  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..180:Ceres Seq. ID 1020785

<400>93  
 Met Thr Ser Ser Tyr Leu Arg Phe Ala Ile Ala Val Val Ala Phe  
 1 5 10 15  
 Leu Ser Ile Thr Thr Ile Thr Thr Ala Arg Pro Cys Lys Thr Phe Leu  
 20 25 30  
 Ile Ser Ser Tyr Ser Leu Ser Ile Thr Pro Glu Asn Pro Asn Leu Glu  
 35 40 45  
 Ser Asp Phe Thr Ser Thr Arg Phe Ile Thr Val Phe Thr Ile Arg Arg  
 50 55 60  
 Leu Asn Pro His Val Val Pro Phe Phe Val Asn Arg Arg His Glu  
 65 70 75 80  
 Lys Pro Gln Ile Gln Ser Asp Arg Ser Leu Pro Leu Ile Ser Asp Asn  
 85 90 95  
 Ile Asn Ser Phe Arg Asp Arg Thr Arg Asp Ile Leu Ser Val Val Val  
 100 105 110  
 Ala Leu Leu Phe Gly Val Gly Cys Gly Ala Leu Thr Ala Thr Met  
 115 120 125  
 Tyr Leu Val Trp Ala Leu Val Val Asn Arg Gln Ser Tyr Asp Phe Glu  
 130 135 140  
 Glu Glu Glu Asp Asp Tyr Glu Asn Asp Glu Ser Asp Ala Ala Ser Leu  
 145 150 155 160

SUBSTITUTE SHEET (RULE 26)



Lys Lys Leu Gly Tyr Val Lys Ile Pro Ala Pro Ala Pro Ala Pro Val  
165 170 175  
Lys Glu Ala Ala  
180

<210> 94  
<211> 765  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..765:Ceres Seq. ID 1021525

<400>94  
acctcaactc atttttcggg tttaaccat tagcacaaaa aatctgagct ccgacacaaa 60  
attgggaac gacgatggc agactcttc agaaactctc ctctcaatc caccgactct 120  
ctccattac cagactctta atcgttgca catcgcaac gtctgtctct tctccgtctc 180  
tgggacgaa gaaagtctct gaccgaatcg tcaagcttct tggatcgat ccgagtgat 240  
acaagcaggaa catcattgga ctctctggcc agactctctt tctgtggctc actcacactg 300  
gtctgataga tccagcatcg catagattgg atgacatcga ggtctactca gcgagtgcg 360  
agtttcagat ccgagaggaa tggctcgaga agctcccgcc gcgtacttac gatgaggagt 420  
atgtctgaa ggtatgctct agactccta ttgtgaacaa gcactcgct ctgtgtgccc 480  
aagttgttt aacgagag cttcaagta tggctgcgc tgcctcgaa gctaacctt 540  
gggatctcc gtaagtttt gtcttgagg atctcntaaa tttgtggctc gtctcaatt 600  
tggctataaa aatcgcttt ggaattcgt gactctgaa gataactcca ttgaaatcg 660  
taatttctg tgcctacat ttaattttg taattgtga aagaactcg ctgttttaa 720  
cgattccctg ctccaagat gggtagatgt tgtttcttg attcc 765

<210> 95  
<211> 159  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..159:Ceres Seq. ID 1021526

<400>95  
Met Ala Thr Thr Leu Gln Lys Lys Leu Ser Ser Gln Ile His Arg Leu Ser  
1 5 10 15  
Pro Phe Thr Arg Ser Leu Ile Val Arg Thr Ser Ala Thr Ser Ala Pro  
20 25  
Ser Pro Ser Leu Gly Ser Lys Lys Val Ser Asp Arg Ile Val Lys Leu  
30 35 40 45  
Ser Ala Ile Asp Pro Asp Gly Tyr Lys Gln Asp Ile Ile Gly Leu Ser  
50 55 60  
Gly Gln Thr Leu Leu Arg Ala Leu Thr His Thr Gly Leu Ile Asp Pro  
65 70 75  
Ala Ser His Arg Leu Asp Asp Ile Glu Ala Cys Ser Ala Glu Cys Glu  
80 85 90  
Val Gln Ile Ala Glu Glu Trp Leu Glu Lys Leu Pro Pro Arg Thr Tyr  
100 105 110  
Asp Glu Glu Tyr Val Leu Lys Arg Ser Ser Arg Ile Leu Asn  
115 120  
Lys His Ser Arg Leu Gly Cys Gln Val Val Leu Thr Gln Glu Leu Gln  
130 140  
Gly Met Val Val Ala Val Pro Glu Ala Lys Pro Trp Asp Ile Pro  
145 155

<210> 96  
<211> 588  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..588:Ceres Seq. ID 1021563

<400>96  
atcttcgaa agtctcaitt ctgatccccc aattctgga ttagggttaa aagaaccatt 60  
tttatctcg tgcgcgaaca acaaatccag atcgaaaaag gaagaagaga tcgaaatggc 120  
tttgagaag gtttaccagtg aaatcagagg gaagaagtg acgagcttc caggctatat 180  
caaatgact ttccaatgg agacctcgg gacctctg agagaggac tcgataacta 240  
caagaaaaa tacattcaga ccagctccgt tgatctatc ctctcatctt gtttaccg 300  
catgcttc tottaccctg tgcctccc taatgagcgt cgcctcttg agcatcaga 360  
gcactaaag gacacaggtg gtactgac tctgggnaat cgttcgctc tggatgat 420  
tttagggggg ttgtgtaga atctttctt gcttgatgt gacgacgat caagaattgt 480  
gtcttatgt tctgtttct tgaatttcc tggataatgt tgacctaaag gaaaccttt 540  
ctttcgatt acactccatg atagtcaata atigaagcat catgatgc 588

<210> 97  
<211> 128  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..128:Ceres Seq. ID 1021564

<400>97  
Ser Ser Glu Ser Leu Ile Ser Arg Ser Pro Ile Arg Gly Leu Gly Leu  
1 5 10 15  
Lys Glu Pro Phe Leu Phe Ser Ser Arg Asn Asn Lys Ser Arg Ser Lys  
20 25 30  
Lys Glu Glu Glu Ile Glu Met Ala Leu Arg Arg Val Tyr Ser Glu Ile  
35 40 45  
Arg Gly Lys Lys Val Thr Glu Leu Pro Gly Tyr Ile Lys Ser Thr Phe  
50 55 60  
Ser Met Glu Thr Val Lys Thr Ser Val Lys Arg Gly Leu Asp Asn Tyr  
65 70 75  
Asn Glu Lys Tyr Ile Gln Thr Ser Ser Val Asp Pro Ile Leu His Ile  
80 85 90  
Cys Phe Tyr Gly Met Ala Phe Ser Tyr Leu Val Ala Leu Pro Asn Glu  
100 105 110  
Arg Arg His Leu Glu His Gln His Ala Lys Glu His Gly Gly His  
115 120 125

<210> 98  
<211> 90  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..90:Ceres Seq. ID 1021565

<400>98  
Met Ala Leu Arg Arg Val Tyr Ser Glu Ile Arg Gly Lys Lys Val Thr  
1 5 10 15  
Glu Leu Pro Gly Tyr Ile Lys Ser Thr Phe Ser Met Glu Thr Val Lys  
20 25 30  
Thr Ser Val Lys Arg Gly Leu Asp Asn Tyr Asn Glu Lys Tyr Ile Gln  
35 40 45  
Thr Ser Ser Val Asp Pro Ile Leu His Ile Cys Phe Tyr Gly Met Ala  
50 55 60  
Phe Ser Tyr Leu Val Ala Leu Pro Asn Glu Arg Arg His Leu Glu His  
65 70 75 80  
Gln Gln His Ala Lys Glu His Gly His  
85 90

<210> 99  
<211> 586  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..586:Ceres Seq. ID 1021576

<400>99  
ttgcctaatt cattttaag actgtataga ggaaacaaaa actgcacaaa caaaataaaa  
60  
aaaacatgc cacaagaaa taagaattt gtagaataca ctaagaaatt ggtagcaact  
120  
atgatgacta catgctacta gtccaattgt cttcgagcca ccaaatctc tgcagctcct  
180  
gtacaagcc tggcaagtgt tcaagccatg agacgaag gaattggagc ttgggtgca  
240  
aagtgtact tcaatgctc atcaacaat ctagaattgg taactgcag gccctgatg  
300  
ttgttcgg ggagattcgg acttgccca tcaagcaata ggaagcgac agctggactt  
360  
agttggagg cagtgactc agttctaca acgggtgacc cggccgggt cactgtgcg  
420  
gacacttgg cttgtggac cgttggtcat atcatcggtg taggattgt ccttggcct  
480  
aaaaacttg gtgtattg agattctcaa agctcttita ttgtattg taaattgtt  
540  
agattttat acaatattc taatgcact gaacgagatc taatgc  
586

<210> 100  
<211> 130  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..130:Ceres Seq. ID 1021577

<400>100  
Met Ala Ser Thr Met Met Thr Thr Leu Pro Gln Phe Asn Gly Leu Arg  
1 5 10 15  
Ala Thr Lys Ile Ser Ala Ala Pro Val Gln Gly Leu Ala Ser Val Gln  
20 25 30  
Pro Met Arg Arg Lys Gly Asn Gly Ala Leu Gly Ala Lys Cys Asp Phe  
35 40 45  
Ile Gly Ser Ser Thr Asn Leu Ile Met Val Thr Ser Thr Thr Leu Met  
50 55 60  
Leu Phe Ala Gly Arg Phe Gly Leu Ala Pro Ser Ala Asn Arg Lys Ala  
65 70 75 80  
Thr Ala Gly Leu Arg Leu Glu Ala Arg Asp Ser Gly Leu Gln Thr Gly  
90 95  
Asp Pro Ala Gly Phe Thr Leu Ala Asp Thr Leu Ala Cys Gly Thr Val  
100 105 110

Gly His Ile Ile Gly Val Gly Val Leu Gly Leu Lys Asn Ile Gly  
115 120 125  
Ala Ile  
130

<210> 101  
<211> 126  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..126:Ceres Seq. ID 1021578

<400>101  
Met Met Thr Thr Leu Pro Gln Phe Asn Gly Leu Arg Ala Thr Lys Ile  
1 5 10 15  
Ser Ala Ala Pro Val Gln Gly Leu Ala Ser Val Gln Pro Met Arg Arg  
20 25 30  
Lys Gly Asn Gly Ala Leu Gly Ala Lys Cys Asp Phe Ile Gly Ser Ser  
35 40 45  
Thr Asn Leu Ile Met Val Thr Ser Thr Thr Leu Met Leu Phe Ala Gly  
50 60  
Arg Phe Gly Leu Ala Pro Ser Ala Asn Arg Lys Ala Thr Ala Gly Leu  
65 70 75 80  
Arg Leu Glu Ala Arg Asp Ser Gly Leu Gln Thr Gly Asp Pro Ala Gly  
90 95  
Phe Thr Leu Ala Asp Thr Leu Ala Cys Gly Thr Val Gly His Ile Ile  
100 105 110  
Gly Val Gly Val Val Leu Gly Leu Lys Asn Ile Gly Ala Ile  
115 120 125

<210> 102  
<211> 125  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..125:Ceres Seq. ID 1021579

<400>102  
Met Thr Thr Leu Pro Gln Phe Asn Gly Leu Arg Ala Thr Lys Ile Ser  
1 5 10 15  
Ala Ala Pro Val Gln Gly Leu Ala Ser Val Gln Pro Met Arg Arg Lys  
20 25 30  
Gly Asn Gly Ala Leu Gly Ala Lys Cys Asp Phe Ile Gly Ser Ser Thr  
35 40 45  
Asn Leu Ile Met Val Thr Ser Thr Thr Leu Met Leu Phe Ala Gly Arg  
50 55 60  
Phe Gly Leu Ala Pro Ser Ala Asn Arg Lys Ala Thr Ala Gly Leu Arg  
65 70 75 80  
Leu Glu Ala Arg Asp Ser Gly Leu Gln Thr Gly Asp Pro Ala Gly Phe  
90 95  
Thr Leu Ala Asp Thr Leu Ala Cys Gly Thr Val Gly His Ile Ile Gly  
100 105 110  
Val Gly Val Val Leu Gly Leu Lys Asn Ile Gly Ala Ile  
115 120 125

&lt;210&gt; 103

&lt;211&gt; 801

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..801:Ceres Seq. ID 1021927

&lt;400&gt;103

attttctct cagaatccat aaaaagagag agagataaat aaagagaaaa ctgaagaagc 60  
 tagaagatg agaaagtaa tgatcatgac aagctatgac acgcgcctc cgggtgtggt 120  
 gcccagaga agtggagaga gactagctc ggaatcagaa ctgcagagac aatgtccgg 180  
 ttactccgg tggctctttg ttgtcagcg ctgtgttga ttgttaagga ctctgagact 240  
 aatgattcg gtcaatttc ttactcaat ctcaacact ttagtactt ggtgaagca 300  
 aatgaatat gtacaggcta ctctctcta tcagcagcca ttacagcat gctcgtct 360  
 tcttcagaa tgcctgtgt ttggccttc ttttgtctg accagctctt gactaacctg 420  
 gtcttgctg ctggagctgt atcagctgag gtctctatac ttgcttacaa tggagactca 480  
 gccattact ggagcagtc atgtatgccc tacggcggtt tctgtcatag agccactgct 540  
 tctgtataa tcacattct ttgtgttgt ttcaaatg ttctctct aatctcctct 600  
 kataagctct tctcgtctt tgatctctt tccattgtg actccgcaa gaatctcgaa 660  
 gtccgtgct tgggaagta gatctccat ttgctctca agtacatcta gttgtgcatg 720  
 ttcaaatgt tgtgtgttt taactttggt caagagaag aatgcttatg tattctctct 780  
 tgttcaatg cttttcttc t 801

&lt;210&gt; 104

&lt;211&gt; 204

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..204:Ceres Seq. ID 1021928

&lt;400&gt;104

Met Glu Lys Ser Asn Asp His Asp Lys Ala Ser His Gly Gly Ser Gly 1  
 Gly Gly Ala Thr Glu Lys Trp Glu Glu Thr Ser Leu Gly Ile Arg Thr 20  
 Ala Glu Thr Met Leu Arg Leu Ala Pro Val Gly Leu Cys Val Ala Ala 35  
 Leu Val Val Met Leu Lys Asp Ser Glu Thr Asn Glu Phe Gly Ser Ile 40  
 Ser Tyr Ser Asn Leu Thr Ala Phe Arg Tyr Leu Val His Ala Asn Gly 50  
 Ile Cys Ala Gly Tyr Ser Leu Leu Ser Ala Ala Ile Ala Ala Met Pro 60  
 Arg Ser Ser Ser Thr Met Pro Arg Val Trp Thr Phe Cys Leu Asp 75  
 Gln Leu Leu Thr Tyr Leu Val Leu Ala Ala Gly Ala Val Ser Ala Glu 80  
 Val Leu Tyr Leu Ala Tyr Asn Gly Asp Ser Ala Ile Thr Trp Ser Asp 95  
 Ala Cys Ser Ser Tyr Gly Gly Phe Cys His Arg Ala Thr Ala Ser Val 100  
 Ile Ile Thr Phe Phe Val Val Cys Phe Tyr Ile Val Leu Ser Leu Ile 110  
 Ser Ser Tyr Lys Leu Phe Thr Arg Phe Asp Pro Ser Ile Val Asp 125  
 130 135 140 145 150 155 160 165 170 175 180 185 190

SUBSTITUTE SHEET (RULE 26)

Ser Ala Lys Asn Leu Glu Val Ala Val Phe Gly Ser 195  
 200

&lt;210&gt; 105

&lt;211&gt; 169

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..169:Ceres Seq. ID 1021929

&lt;400&gt;105

Met Leu Arg Leu Ala Pro Val Gly Leu Cys Val Ala Ala Leu Val Val 1  
 Met Leu Lys Asp Ser Glu Thr Asn Glu Phe Gly Ser Ile Ser Tyr Ser 10  
 Asn Leu Thr Ala Phe Arg Tyr Leu Val His Ala Asn Gly Ile Cys Ala 20  
 Gly Tyr Ser Leu Leu Ser Ala Ala Ile Ala Ala Met Pro Arg Ser Ser 25  
 Ser Thr Met Pro Arg Val Trp Thr Phe Cys Leu Asp Gln Leu Leu 30  
 Thr Tyr Leu Val Leu Ala Ala Gly Ala Val Ser Ala Glu Val Leu Tyr 35  
 Leu Ala Tyr Asn Gly Asp Ser Ala Ile Thr Trp Ser Asp Ala Cys Ser 40  
 Ser Tyr Gly Gly Phe Cys His Arg Ala Thr Ala Ser Val Ile Ile Thr 45  
 Phe Phe Val Val Cys Phe Tyr Ile Val Leu Ser Leu Ile Ser Ser Tyr 50  
 Lys Leu Phe Thr Arg Phe Asp Pro Pro Ser Ile Val Asp Ser Ala Lys 55  
 Asn Leu Glu Val Ala Val Phe Gly Ser 60  
 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165

&lt;210&gt; 106

&lt;211&gt; 153

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..153:Ceres Seq. ID 1021930

&lt;400&gt;106

Met Leu Lys Asp Ser Glu Thr Asn Glu Phe Gly Ser Ile Ser Tyr Ser 1  
 Asn Leu Thr Ala Phe Arg Tyr Leu Val His Ala Asn Gly Ile Cys Ala 10  
 Gly Tyr Ser Leu Leu Ser Ala Ala Ile Ala Ala Met Pro Arg Ser Ser 20  
 Ser Thr Met Pro Arg Val Trp Thr Phe Cys Leu Asp Gln Leu Leu 25  
 Thr Tyr Leu Val Leu Ala Ala Gly Ala Val Ser Ala Glu Val Leu Tyr 30  
 Leu Ala Tyr Asn Gly Asp Ser Ala Ile Thr Trp Ser Asp Ala Cys Ser 35  
 Ser Tyr Gly Gly Phe Cys His Arg Ala Thr Ala Ser Val Ile Ile Thr 40  
 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165

SUBSTITUTE SHEET (RULE 26)

100  
Phe Phe Val Val Cys Phe Tyr Ile Val Ser Ser Tyr  
115  
Lys Leu Phe Thr Arg Phe Asp Pro Pro Ser Ile Val Asp Ser Ala Lys  
130  
Asn Leu Glu Val Ala Val Phe Gly Ser  
145

<210> 107  
<211> 602  
<212> DNA  
<213> Arabidopsis thaliana  
<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..602:Ceres Seq. ID 1021945

<400>107  
atcgagcgat ggatttggca gaactgtgg cgattttcgg acccgattc tccggccgcg  
ttttcgaac cgggtgtgg ggtcagcc gtcgttgcga gtccatcca agttcccttc  
gtccattacc ttccggtcat attcgtctt ctcgagctt tgaatttcaa ttggtcaga  
aaagaaga ttgatttacc tcttaccac gaagcgagtt ggaattgaa gctgtgctt  
ttcatagct atgtcgtac attgtttcc ttacgtctt cgttggctt gctgattcaa  
gattcgttg tgaagactgg gcttcaact tggactgtg tggctgtgt atttcaatgt  
gtattgtat tgaatagtg gctaatgat tggacatgc actcagagta gcgcacagc  
acatactca tctccgactc agattcaac tbtgacaact tacaagtcc tctcgtttt  
ctactatga aattcctgtg ttctcgttg tcaacttggc atttggcg agattcctt  
gtgtcatcag tgtgtgttc tgttaaaact ttgtgtgat gttattgaaa atactccaaa  
cc

<210> 109  
<211> 136  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..136:Ceres Seq. ID 1021946

<400>108  
Ile Glu Arg Trp Ile Trp Gln Asn Cys Gly Arg Phe Ser Asp Pro Asp  
1  
Ser Pro Ala Pro Phe Ser Glu Pro Gly Gly Val Asp Ala Val Val  
Cys Ser Ser Ile Gln Val Pro Phe Val His Tyr Leu Pro Gly Ile Phe  
Ala Ser Leu Gly Ala Leu Met Phe Asn Cys Val Arg Lys Glu Asp Ile  
Asp Tyr Ser Pro Tyr Asp Glu Gly Trp Arg Leu Lys Leu Trp Leu  
Phe Ile Ala Tyr Val Val Ala Phe Val Ser Leu Ala Ala Ser Val Gly  
Leu Leu Ile Gln Asp Ser Val Val Lys Thr Gly Pro Ser Thr Trp Thr  
Gly Val Ala Gly Val Phe Gln Cys Val Phe Val Leu Ile Ser Gly Leu  
Met Tyr Trp Thr Ser His Ser Glu

<210> 109  
<211> 82  
<212> PRT  
<213> Arabidopsis thaliana  
<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..82:Ceres Seq. ID 1021947

<400>109  
Met Phe Asn Cys Val Arg Lys Glu Asp Ile Asp Tyr Ser Pro Tyr Asp  
1  
Glu Gly Glu Trp Arg Leu Lys Leu Trp Leu Phe Ile Ala Tyr Val Val  
Ala Phe Val Ser Leu Ala Ala Ser Val Gly Leu Leu Ile Gln Asp Ser  
Val Val Lys Thr Gly Pro Ser Thr Trp Thr Gly Val Ala Gly Val Phe  
Gln Cys Val Phe Val Leu Ile Ser Gly Leu Met Tyr Trp Thr Ser His  
Ser Glu

<210> 110  
<211> 693  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..693:Ceres Seq. ID 1022170

<400>110  
gagtcgctt tcttctctc cgcattttt tctttctag gttcagagt taaagagaa  
ggtttcaat taggttttg tagagaaaa gatgagcga agtttggaa taccgtgaa  
gctttcac gaggctcag gtcatacgt gcggtggag taaagagcg gcagattta  
cagagaaat atgattgagt gtgagatata ctggaactgt cagctcgag atattacta  
taccgcaag gatgttaag tatcacagt tgagcatgtc ttcatcgag gcagtaagt  
caggttatg gtcataccag acatttcaa acatgtcca atgtcaagc ggttagatc  
tgaatcaag ggaagagct catcactggg tgttgcaga gttagagctg caatgcgag  
gaacccgct gccggcccg ggcgtggaac tggaggaag ggagcgtac caccgtgag  
gagatgattg atttaccgt caacatgaa gctttgattc tggatagtc ttctgtgat  
atccagtgga aagcgcctt tcttggttct gtcattaatg ctttagagaa cactgaaaa  
cnaacaaga gagagatata tccagatgca ctgtagtctc tgtatttgg gttaccaaga  
tgttgaact attgacaagt ttctgtgtg atc

<210> 111  
<211> 131  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..131:Ceres Seq. ID 1022171

<400>111  
Met Ser Arg Ser Leu Gly Ile Pro Val Lys Leu Leu His Glu Ala Ser

Gly His Ile Val Thr Val Glu Leu Lys Ser Gly Glu Leu Tyr Arg Gly  
 20 25 30  
 Ser Met Ile Glu Cys Glu Asp Asn Trp Asn Cys Gln Leu Glu Asp Ile  
 35 40 45  
 Thr Tyr Thr Ala Lys Asp Gly Lys Val Ser Gln Leu Glu His Val Phe  
 50 55 60  
 Ile Arg Gly Ser Lys Val Arg Phe Met Val Ile Pro Asp Ile Leu Lys  
 65 70 75 80  
 His Ala Pro Met Phe Lys Arg Leu Asp Ala Arg Ile Lys Gly Lys Ser  
 85 90 95  
 Ser Ser Leu Gly Val Gly Arg Gly Arg Ala Ala Met Arg Gly Lys Pro  
 100 105 110  
 Ala Ala Gly Pro Gly Arg Gly Thr Gly Arg Gly Ala Val Pro Pro  
 115 120 125  
 Val Arg Arg  
 130

&lt;210&gt; 112

&lt;211&gt; 98

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..98:Ceres Seq. ID 1022172

&lt;400&gt;112

Met Ile Glu Cys Glu Asp Asn Trp Asn Cys Gln Leu Glu Asp Ile Thr  
 1 5 10 15  
 Tyr Thr Ala Lys Asp Gly Lys Val Ser Gln Leu Glu His Val Phe Ile  
 20 25 30  
 Arg Gly Ser Lys Val Arg Phe Met Val Ile Pro Asp Ile Leu Lys His  
 35 40 45  
 Ala Pro Met Phe Lys Arg Leu Asp Ala Arg Ile Lys Gly Lys Ser Ser  
 50 55 60  
 Ser Leu Gly Val Gly Arg Gly Arg Ala Ala Met Arg Gly Lys Pro Ala  
 65 70 75 80  
 Ala Gly Pro Gly Arg Gly Thr Gly Gly Arg Gly Ala Val Pro Pro Val  
 85 90 95  
 Arg Arg

&lt;210&gt; 113

&lt;211&gt; 65

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..65:Ceres Seq. ID 1022173

&lt;400&gt;113

Glu Ser Leu Phe Phe Pro Ser Ala Ile Phe Phe Leu Glu Gly Cys Arg  
 1 5 10 15  
 Val Lys Gly Glu Gly Phe Gln Leu Gly Phe Cys Arg Glu Lys Asp Glu  
 20 25 30  
 Pro Lys Phe Gly Asn Thr Gly Glu Ala Ser Ser Arg Gly Leu Arg Ser  
 35 40 45  
 Tyr Arg Asp Gly Gly Ala Lys Glu Arg Arg Ala Leu Gln Arg Lys Tyr

50 55 60

Asp

65

&lt;210&gt; 114

&lt;211&gt; 706

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..706:Ceres Seq. ID 1022554

&lt;400&gt;114

agacacaaa agcggttggt gtaagatccc aaactcacg attcccaaat aatagtaata  
 ctctccctct tctcaactct caccagtcac cagcagatca tggagatgg gcggagacgg  
 aaaaatttcc acctgtccg aggtttctca gcacagtagc gccaaaggatt gtggatcgt  
 catcagcggc aaggtttatg atgtgacaaa gtctctggat gatcctctg gtggatgta  
 ggtaattctg acctctacag ggaagatgc gaccgatgat ttcgaggatg tgggacatag  
 ttgcactcgc aaagccatgc tagatagta ctatgtgggt gatattgaca cagctactgt  
 gccggttaaa gctaagtttg tgcctctac gtgcacgaaa gccgtggcta ctcaggtaa  
 gagctcggat ttgtttatta agctctctca gtctcttgtt ccactctcaa tcttaggctt  
 ggctctggc attcgttact acactaagac caaggtctct tcttctgaa agattgaatt  
 gtggcctaag tgaattgctt gtgtacgacg tggitaacct ctgtctctac cctctgttgt  
 ctgtgaagac atcattatat tctctattgc aatcaactga tagaactttg attttaagc  
 catctacctt tgtgtcttta tatcaatcaa atccatttac aatatac

&lt;210&gt; 115

&lt;211&gt; 140

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..140:Ceres Seq. ID 1022555

&lt;400&gt;115

Met Gly Gly Asp Gly Lys Val Phe Thr Leu Ser Glu Val Ser Gln His  
 1 5 10 15  
 Ser Ser Ala Lys Asp Cys Trp Ile Val Ile Asp Gly Lys Val Tyr Asp  
 20 25 30  
 Val Thr Lys Phe Leu Asp Asp His Pro Gly Asp Glu Val Ile Leu  
 35 40 45  
 Thr Ser Thr Gly Lys Asp Ala Thr Asp Asp Phe Glu Asp Val Gly His  
 50 55 60  
 Ser Ser Thr Ala Lys Ala Met Leu Asp Glu Tyr Tyr Val Gly Asp Ile  
 65 70 75 80  
 Asp Thr Ala Thr Val Pro Val Lys Ala Lys Phe Val Pro Pro Thr Ser  
 85 90 95  
 Thr Lys Ala Val Ala Thr Gln Asp Lys Ser Ser Asp Phe Val Ile Lys  
 100 105 110  
 Leu Leu Gln Phe Leu Val Pro Leu Leu Ile Leu Gly Leu Ala Phe Gly  
 115 120 125  
 Ile Arg Tyr Tyr Thr Lys Thr Lys Ala Pro Ser Ser  
 130 135 140

&lt;210&gt; 116

&lt;211&gt; 70

&lt;212&gt; PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..70:Ceres Seq. ID 1022556

<400>116

Met Leu Asp Glu Tyr Tyr Val Gly Asp Ile Asp Thr Ala Thr Val Pro  
1 5 10 15  
Val Lys Ala Lys Phe Val Pro Pro Thr Thr Lys Ala Val Ala Thr  
20 25 30  
Gln Asp Lys Ser Ser Asp Phe Val Ile Lys Leu Leu Gln Phe Leu Val  
35 40 45  
Pro Leu Leu Ile Leu Gly Leu Ala Phe Gly Ile Arg Tyr Tyr Thr Lys  
50 55 60  
Thr Lys Ala Pro Ser Ser  
65 70

<210> 117

<211> 722

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..722:Ceres Seq. ID 1022594

<400>117

aactcatea ttctgacttt tctcaagaa gagaaatcaa atccaccaagt catcgagcg  
60  
atccagaga agaagaaac catggcgaa gaagagaata gctgtatgaa gcaagtgtg  
120  
gtggcgattg atgaagcga atgtagttaa cgcgtcttc aatgagcgt cgtgtatctc  
180  
aaagatgacc tgcgcgatic cgcacatcac ctcttcactg ctcaactca ttctgatctc  
240  
agcgtgct atgtctctc ttatggccc gctccgattg agcttataa ctctattcaa  
300  
gagatcata agaagctg attgaatgt cttgatgaag gaaccaaat ttgtgctgag  
360  
actgggita ctcaagaa ggtgttgaa ttggaaatc ctcaagagc gatagttaa  
420  
gctgctaga agcttggtg tgatatgt gtgtttgaa gccatggtaa aggagcata  
480  
caaggactt tcttggaag tgttagaat tactgttta acaatgctaa gtgccagtt  
540  
cttgttga gaacaaagc ttgaagact ctctcatga aagagttcga caacctggc  
600  
ttgatgtg tgtgtgttaa acatatgt ataatgctc tcttgggta ctattgtt  
660  
aagaacctt gtgagttaac ataatttatt ttgttgata atgaagagc attgccgtg  
720  
tc 722

<210> 118

<211> 187

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..187:Ceres Seq. ID 1022595

<400>118

Asn Val Ile Ile Leu Thr Phe Leu Lys Glu Glu Lys Ser Asn His Gln  
1 5 10 15  
Val Ile Gly Ala Ile Thr Glu Lys Lys Thr Met Ala Glu Glu  
20 25 30  
Lys Ser Val Met Lys Gln Val Met Val Ala Ile Asp Glu Ser Glu Cys  
35 40 45

Ser Lys Arg Ala Leu Gln Tyr Thr Leu Val Tyr Leu Lys Asp Ser Leu  
50 55 60  
Ala Asp Ser Asp Ile Ile Leu Phe Thr Ala Gln Pro His Leu Asp Leu  
65 70 75 80  
Ser Cys Val Tyr Ala Ser Ser Tyr Gly Ala Ala Pro Ile Glu Leu Ile  
90 95  
Asn Ser Leu Gln Glu Ser His Lys Asn Ala Gly Leu Asn Arg Leu Asp  
100 105 110  
Glu Gly Thr Lys Ile Cys Ala Glu Thr Gly Val Thr Pro Arg Lys Val  
115 120 125  
Leu Glu Phe Gly Asn Pro Lys Glu Ala Ile Cys Glu Ala Ala Glu Lys  
130 135 140  
Leu Gly Val Asp Met Leu Val Val Gly Ser His Gly Lys Gly Ala Leu  
145 150 155 160  
Gln Arg Thr Phe Leu Gly Ser Val Ser Asn Tyr Cys Val Asn Asn Ala  
170 175  
Lys Cys Pro Val Leu Val Val Arg Thr Lys Ala  
180 185

<210> 119

<211> 160

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..160:Ceres Seq. ID 1022596

<400>119

Met Ala Glu Glu Glu Lys Ser Val Met Lys Gln Val Met Val Ala Ile  
1 5 10 15  
Asp Glu Ser Glu Cys Ser Lys Arg Ala Leu Gln Tyr Thr Leu Val Tyr  
20 25 30  
Leu Lys Asp Ser Leu Ala Asp Ser Asp Ile Ile Leu Phe Thr Ala Gln  
35 40 45  
Pro His Leu Asp Leu Ser Cys Val Tyr Ala Ser Ser Tyr Gly Ala Ala  
50 55 60  
Pro Ile Glu Leu Ile Asn Ser Leu Gln Glu Ser His Lys Asn Ala Gly  
65 70 75 80  
Leu Asn Arg Leu Asp Glu Gly Thr Lys Ile Cys Ala Glu Thr Gly Val  
85 90 95  
Thr Pro Arg Lys Val Leu Glu Phe Gly Asn Pro Lys Glu Ala Ile Cys  
100 105 110  
Glu Ala Ala Glu Lys Leu Gly Val Asp Met Leu Val Val Gly Ser His  
115 120 125  
Gly Lys Gly Ala Leu Gln Arg Thr Phe Leu Gly Ser Val Ser Asn Tyr  
130 135 140  
Cys Val Asn Asn Ala Lys Cys Pro Val Leu Val Val Arg Thr Lys Ala  
145 150 155 160

<210> 120

<211> 152

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..152:Ceres Seq. ID 1022597

<400>120  
Met Lys Gln Val Met Val Ala Ile Asp Glu Ser Glu Cys Ser Lys Arg  
1 5 10 15  
Ala Leu Gln Trp Thr Leu Val Tyr Leu Lys Asp Ser Leu Ala Asp Ser  
20 25 30  
Asp Ile Ile Leu Phe Thr Ala Gln Pro His Leu Asp Leu Ser Cys Val  
35 40 45  
Tyr Ala Ser Ser Tyr Gly Ala Ala Pro Ile Glu Leu Ile Asn Ser Leu  
50 55 60  
Gln Glu Ser His Lys Asn Ala Gly Leu Asn Arg Leu Asp Glu Gly Thr  
65 70 75 80  
Lys Ile Cys Ala Glu Thr Gly Val Thr Pro Arg Lys Val Leu Glu Phe  
85 90 95  
Gly Asn Pro Lys Glu Ala Ile Cys Glu Ala Ala Glu Lys Leu Gly Val  
100 105 110  
Asp Met Leu Val Val Gly Ser His Gly Lys Gly Ala Leu Gln Arg Thr  
115 120 125  
Phe Leu Gly Ser Val Ser Asn Tyr Cys Val Asn Asn Ala Lys Cys Pro  
130 135 140  
Val Leu Val Val Arg Thr Lys Ala  
145 150

<210> 121  
<211> 610  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..610:Ceres Seq. ID 1022621

<400>121  
aatttgggt ggtatattac tggtagaga gttcaaatc caatctttct ctctctct  
atctttatc tctcaatc ttaaatgct tctctatc ccatctctcgc cctctcttc  
accgtgttt caccgggta attgatccg gcatcaatt aagcgttag taggttgt  
tgaagaagc aaggacnaca cgcacatgg tggattctg gagagcgag ctatagcgg  
tggtttagt tgcagccgg ttttggatg gcatctac actttgaaga cccgggatg  
tggcttact cccgagccgg cgggtttaat cggagcgtt gaggtgtga gctacttgt  
360 ggtggtggc atcgiggct ggtcttga cactaaaca aaactgggt caggtctgc  
420 caatctgt ggtttgta ttcagctctt ggaataagg tgggtccag gtcacttcc  
480 aatgggcca ttgtttgtg tgggggggt caggggtta tctatctgt cgtttcagc  
540 tadtgaccag tgttttggt aaccacgtg aatgctgat aaataaagtg tcaattaat  
600 aattgtttc 610

<210> 122  
<211> 158  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..158:Ceres Seq. ID 1022622

<400>122  
Met Leu Leu Leu Ser Pro Ile Ser Ala Ser Leu Pro Pro Ser Phe His  
1 5 10 15  
Arg Gly Asn Leu Ile Arg Arg Ser Ile Lys Pro Leu Gly Arg Val Val  
20 25 30  
Ala Lys Ala Lys Asp Asn Thr Asp Thr Gly Gly Phe Leu Glu Thr Ala

35 40 45  
Ala Ile Ala Gly Gly Leu Val Ser Thr Pro Val Ile Gly Trp Ser Leu  
50 55 60  
Tyr Thr Leu Lys Thr Thr Gly Cys Gly Leu Pro Gly Pro Ala Gly  
65 70 75 80  
Leu Ile Gly Ala Leu Glu Gly Val Ser Tyr Leu Val Val Gly Ile  
85 90 95  
Val Gly Trp Ser Leu Tyr Thr Lys Thr Lys Thr Gly Ser Gly Leu Pro  
100 105 110  
Asn Gly Pro Phe Gly Leu Leu Gly Ala Val Glu Gly Leu Ser Tyr Leu  
115 120 125  
Ser Val Leu Ala Ile Leu Val Val Phe Gly Ile Gln Phe Leu Asp Asn  
130 135 140  
Gly Ser Val Pro Gly Pro Leu Pro Ser Asp Gln Cys Phe Gly  
145 150 155

<210> 123  
<211> 872  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..872:Ceres Seq. ID 1024375

<400>123  
atcgaacatt aaacctgat ttactgaa ttgtgtgag aaggagaagc cgaacacat  
ttttgtcaa tctaggtgtc ttctgttac ttcagattt cgcctcgtt ttcactaga  
120 tctgtgaaga tgtctccgc tcagaacaag atcaataagg acagaatgc tgaaccaaca  
180 gaatgcgagg agcaagtgc tcaggtttg ttgatttgg agaacactaa ccaggatgg  
240 aaagcgagt tgaagatct ctaatcaac caagctgttc acatgatat ctctgaaac  
300 cgcgaactg ttgtgattt cgttcattc agattgaga aagcttccg caagattoat  
360 cccgtctcg tcagagatc ttgagaag ttgagtgga aggatgttat cttgttacc  
420 acaagaaga tcatgctcc cccaagaag ggtgctgtg ttcadagcc acgtaacaga  
480 actttacct cagtcatga agttatggt gaagatgtt cttccccc tgagattgt  
540 ggaagcgta ctgcctacg tcttgatgt tccaagatca tgaagttct ttggatgcc  
600 agggaaaga acacacaga gtacacgtc gacatctgg tgggtgtga cgttaacct  
660 actggcaag atgttgttt tggatccca cgaagctt gaaagaagt gatgaagaac  
720 catcagata gtgaagaga gcttttgtt atgtttgtg gtattttaga tgaaggaac  
780 tctcttatt cagttccgtg ttcacatct ttaatttct atttacaatg actactttt  
840 gtgttttca atttgaac ctctatgatt tc 872

<210> 124  
<211> 190  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..190:Ceres Seq. ID 1024376

<400>124  
Met Phe Ser Ala Gln Asn Lys Ile Asn Lys Asp Lys Asn Ala Glu Pro  
1 5 10 15  
Thr Glu Cys Glu Glu Gln Val Ala Gln Ala Leu Phe Asp Leu Glu Asn  
20 25 30  
Thr Asn Gln Glu Leu Lys Ser Glu Leu Lys Asp Leu Tyr Ile Asn Gln  
35 40 45  
Ala Val His Met Asp Ile Ser Gly Asn Arg Lys Ala Val Val Ile Tyr

50 Val Pro Phe Arg Leu Arg Lys Ala Phe Arg Lys Ile His Pro Arg Leu 60  
 65 Val Arg Glu Leu Glu Lys Lys Phe Ser Gly Lys Asp Val Ile Phe Val 75  
 80 Thr Thr Arg Arg Ile Met Arg Pro Pro Lys Lys Gly Ala Ala Val Gln 90  
 95 Arg Pro Arg Asn Arg Thr Leu Thr Ser Val His Glu Ala Met Leu Glu 110  
 115 Asp Val Ala Phe Pro Ala Glu Ile Val Gly Lys Arg Thr Arg Tyr Arg 125  
 130 Leu Asp Gly Ser Lys Ile Met Lys Val Phe Leu Asp Ala Lys Glu Lys 140  
 145 Asn Asn Thr Glu Tyr Lys Leu Glu Thr Met Val Gly Val Tyr Arg Lys 155  
 160 Leu Thr Gly Lys Asp Val Val Phe Glu Tyr Pro Val Glu Ala 175  
 180 185 190

<210> 125  
 <211> 139  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..139:Ceres Seq. ID 1024377

<400>125  
 Met Asp Ile Ser Gly Asn Arg Lys Ala Val Val Ile Tyr Val Pro Phe 15  
 1 Arg Leu Arg Lys Ala Phe Arg Lys Ile His Pro Arg Leu Val Arg Glu 20  
 25 Leu Glu Lys Lys Phe Ser Gly Lys Asp Val Ile Phe Val Thr Thr Arg 30  
 35 Arg Ile Met Arg Pro Pro Lys Lys Gly Ala Ala Val Gln Arg Pro Arg 45  
 50 Asn Arg Thr Leu Thr Ser Val His Glu Ala Met Leu Glu Asp Val Ala 60  
 65 Phe Pro Ala Glu Ile Val Gly Lys Arg Thr Arg Tyr Arg Leu Asp Gly 75  
 80 Ser Lys Ile Met Lys Val Phe Leu Asp Ala Lys Glu Lys Asn Asn Thr 90  
 100 Glu Tyr Lys Leu Glu Thr Met Val Gly Val Tyr Arg Lys Leu Thr Gly 110  
 115 Lys Asp Val Val Phe Glu Tyr Pro Val Glu Ala 125  
 130 135

<210> 126  
 <211> 545  
 <212> DNA  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..545:Ceres Seq. ID 1024535

<400>126  
 aaaaattgtt aagttcgtgt ccgcgagatc agagtttgtg tccggttata cgtatcgtcg 60  
 gtacaaaac cctccagaa gtgtttcagg tgggaaggta tctctgtcta caagatgtcc 120

ggcgggaag aagccaccgt gagggagcca ctgacttga ttaggtgag tctcgacgag 180  
 agaatctatg tcaagctccg gtcagaccgc gaacttcgcg gcaagttcca cgcgtttgat 240  
 cagcattga atatgattct ggtgatgtt gaagaaacta tcactacagt aaaaatcagt 300  
 gacgagcat atgaagagat tgttcggact acaagcggg cgaattgagt ttatttcgtg 360  
 agaggagatg gagtgaatt ggtctcca ccgtcaggga cagcagcttg agttcaact 420  
 caaatcttt ggtctcttt ttctgaaggg gaatttga ctattctgt atttagtcaac 480  
 ttctgtctt aactaaaca agcttgaga agctagtgtg cttaaatctt caatgctttt 540  
 cagac 545

<210> 127  
 <211> 98  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..98:Ceres Seq. ID 1024536

<400>127  
 Met Ser Gly Glu Glu Glu Ala Thr Val Arg Glu Pro Leu Asp Leu Ile 15  
 1 Arg Leu Ser Leu Asp Glu Arg Ile Tyr Val Lys Leu Arg Ser Asp Arg 20  
 25 Glu Leu Arg Gly Lys Leu His Ala Phe Asp Gln His Leu Asn Met Ile 30  
 35 Leu Gly Asp Val Glu Glu Thr Thr Thr Val Glu Ile Asp Asp Glu 45  
 50 Thr Tyr Glu Glu Ile Val Arg Thr Thr Lys Arg Thr Thr Ile Glu Phe Leu 55  
 60 Phe Val Arg Gly Asp Gly Val Ile Leu Val Ser Pro Pro Leu Arg Thr 65  
 70 75 80 85 90 95  
 Ala Ala

<210> 128  
 <211> 53  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..53:Ceres Seq. ID 1024537

<400>128  
 Lys Leu Leu Ser Ser Cys Pro Arg Asp Gln Ser Leu Cys Pro Val Ile 15  
 1 Arg Phe Val Gly Thr Lys Thr Leu Arg Glu Val Phe Gln Val Gly Arg 20  
 25 Tyr Leu Cys Leu His Asp Val Arg Arg Gly Arg Ser His Arg Glu Gly 30  
 35 Ala Thr Arg Ser Asp 40 45 50

<210> 129  
 <211> 52  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>



<223> any n or Xaa = unknown

<223> LOCATION: 1..52:Ceres Seq. ID 1024538

<400>129  
Met Ile Leu Gly Asp Val Glu Glu Thr Ile Thr Thr Val Glu Ile Asp  
1 5 10 15  
Asp Glu Thr Tyr Glu Glu Ile Val Arg Thr Thr Lys Arg Thr Ile Glu  
20 25 30  
Phe Leu Phe Val Arg Gly Asp Gly Val Ile Leu Val Ser Pro Pro Leu  
35 40 45  
Arg Thr Ala Ala  
50

<210> 130

<211> 797

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..797:Ceres Seq. ID 1025683

<400>130  
aacccitaaa ctcggttaac ctaaacctca agccgtctgc tccaaagttc cccctcagag  
120 acgaaaaat gaagaccatt ttgtcctcgg agactatgga catcccyac ggctgcgcca  
180 ttaagtaaa cgcgaagtg attgaggtcg aaggtccacg aggtaaactc actcgtgaat  
240 tcaagcatct gaatcagat ttcagcttga ttaaaagcca agtcaactga aaacgtcagc  
300 ttaagattga ttcttggttt ggtctcctga agcaaatgc ttcgattaga actcgtttaa  
360 gcaatgtga taatccatt gcgggttt ccaaggttt tctttataga atgagatttc  
420 tgaatgcta ttttctatc aatgcttcta ttgatggtaa caataaagat attgagattc  
480 gtaactctc tggtagaag aaggtgaga aggttgagat gttgagtggt attaagattg  
540 ttcgattga gaaggttaag gacgagatta tcttgaggg aaagatcatt gacgttggtt  
600 cagcgtcttg tggtttgatc aatcagaat gtcattgtaa gaagaagat atcgggaat  
660 tcttgatgg taactatgtg agcgagaag gcaagatcgc agtcaggaa tgaattcgc  
720 tatgaagt tctatattag attttatag ggaagcagc gatattatcg tagcttttg  
780 tttcacatc ttttatgga ttgttagttt tgttttctc ttgtaacat tgaatgatc  
797 aaattcatta tccgttc

<210> 131

<211> 216

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..216:Ceres Seq. ID 1025684

<400>131  
Pro Leu Asn Ser Val Thr Leu Asn Leu Lys Lys Pro Ser Ser Pro Lys Phe  
1 5 10 15  
Pro Leu Arg Asp Glu Lys Met Lys Thr Ile Leu Ser Ser Glu Thr Met  
20 25 30  
Asp Ile Pro Asp Gly Val Ala Ile Lys Val Asn Ala Lys Val Ile Glu  
35 40 45  
Val Glu Gly Pro Arg Gly Lys Leu Thr Arg Asp Phe Lys Lys Leu Asn  
50 55 60  
Leu Asp Phe Gln Leu Ile Lys Asp Gln Val Thr Gly Lys Arg Gln Leu  
65 70 75 80

Lys Ile Asp Ser Trp Phe Gly Ser Arg Lys Thr Ser Ala Ser Ile Arg  
85 90 95  
Thr Ala Leu Ser His Val Asp Asn Leu Ile Ala Gly Val Thr Gln Gly  
100 105 110  
Phe Leu Tyr Arg Met Arg Phe Val Tyr Ala His Phe Pro Ile Asn Ala  
115 120 125  
Ser Ile Asp Gly Asn Asn Lys Ser Ile Glu Ile Arg Asn Phe Leu Gly  
130 135 140  
Glu Lys Lys Val Arg Lys Val Glu Met Leu Asp Gly Val Lys Ile Val  
145 150 155  
Arg Ser Glu Lys Val Lys Asp Glu Ile Ile Leu Glu Gly Asn Asp Ile  
165 170 175  
Glu Leu Val Ser Arg Ser Cys Ala Leu Ile Asn Gln Lys Cys His Val  
180 185 190  
Lys Lys Lys Asp Ile Arg Lys Phe Leu Asp Gly Ile Tyr Val Ser Glu  
195 200 205  
Lys Gly Lys Ile Ala Val Glu Glu  
210 215

<210> 132

<211> 194

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..194:Ceres Seq. ID 1025685

<400>132  
Met Lys Thr Ile Leu Ser Ser Glu Thr Met Asp Ile Pro Asp Gly Val  
1 5 10 15  
Ala Ile Lys Val Asn Ala Lys Val Ile Glu Val Glu Gly Pro Arg Gly  
20 25 30  
Lys Leu Thr Arg Asp Phe Lys His Leu Asn Leu Asp Phe Gln Leu Ile  
35 40 45  
Lys Asp Gln Val Thr Gly Lys Arg Gln Leu Lys Ile Asp Ser Trp Phe  
50 55 60  
Gly Ser Arg Lys Thr Ser Ala Ser Ile Arg Thr Ala Leu Ser His Val  
65 70 75 80  
Asp Asn Leu Ile Ala Gly Val Thr Gln Gly Phe Leu Tyr Arg Met Arg  
85 90 95  
Phe Val Tyr Ala His Phe Pro Ile Asn Ala Ser Ile Asp Gly Asn Asn  
100 105 110  
Lys Ser Ile Glu Ile Arg Asn Phe Leu Gly Glu Lys Lys Val Arg Lys  
115 120 125  
Val Glu Met Leu Asp Gly Val Lys Ile Val Arg Ser Glu Lys Val Lys  
130 135 140  
Asp Glu Ile Ile Leu Glu Gly Asn Asp Ile Glu Leu Val Ser Arg Ser  
145 150 155  
Cys Ala Leu Ile Asn Gln Lys Cys His Val Lys Lys Lys Asp Ile Arg  
165 170 175  
Lys Phe Leu Asp Gly Ile Tyr Val Ser Glu Lys Gly Lys Ile Ala Val  
180 185 190  
Glu Glu

<210> 133

<211> 185

<212> PRT

<213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..185:Ceres Seq. ID 1025686

<400>133  
 Met Asp Ile Pro Asp Gly Val Ala Ile Lys Val Asn Ala Lys Val Ile  
 1 5 10 15  
 Glu Val Glu Gly Pro Arg Gly Lys Leu Thr Arg Asp Phe Lys His Leu  
 20 25 30  
 Asn Leu Asp Phe Gln Leu Ile Lys Asp Gln Val Thr Gly Lys Arg Gln  
 35 40 45  
 Leu Lys Ile Asp Ser Trp Phe Gly Ser Arg Lys Thr Ser Ala Ser Ile  
 50 55 60  
 Arg Thr Ala Leu Ser His Val Asp Asn Leu Ile Ala Gly Val Thr Gln  
 65 70 75  
 Gly Phe Leu Tyr Arg Met Arg Phe Val Tyr Ala His Phe Pro Ile Asn  
 80 85 90 95  
 Ala Ser Ile Asp Gly Asn Asn Lys Ser Ile Glu Ile Arg Asn Phe Leu  
 100 105 110  
 Gly Glu Lys Lys Val Arg Lys Val Glu Met Leu Asp Gly Val Lys Ile  
 115 120 125  
 Val Arg Ser Glu Lys Val Lys Asp Glu Ile Ile Leu Glu Gly Asn Asp  
 130 135 140  
 Ile Glu Leu Val Ser Arg Ser Cys Ala Leu Ile Asn Gln Lys Cys His  
 145 150 155  
 Val Lys Lys Lys Asp Ile Arg Lys Phe Leu Asp Gly Ile Tyr Val Ser  
 160 165 170 175  
 Glu Lys Gly Lys Ile Ala Val Glu Glu  
 180 185

<210> 134

<211> 555

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..555:Ceres Seq. ID 1027152

<400>134  
 gttttgca tctgtctcct cttctctcc accgcacaga tccaagacat ttgtattaca 60  
 cccatgcc ggcgaataat gtacggaaa agaaacgaa gaagtctcat gaggtatca 120  
 acagttagtt ggtcttgttt atgaagtg gtaaatcac tctgtgtac agtctgttc 180  
 tcaatccct cgtggctcc aaaggaaac tgattctaat ctccaccaat tgcacccgt 240  
 tgaagatc agagattgag tactacgaa tgcctgctaa agtttgttc caccattaca 300  
 atggaaataa cttgtattgg ggaacagctt gcgaaagata ctccctgttt tctgtctga 360  
 gatttgtga tctgtgtgat tctgaatta tcaagtcat tctgtgtgac cagtgtaca 420  
 atttccagat ttatttttc ccagacttt taccgtctt tgtattcag tctctcaaa 480  
 tttatgac aatagacatt gtgttcttc ttgtccaagt ttctttatat tgaagacgg 540  
 aaagaaatt tcttt 555

<210> 135

<211> 137

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..137:Ceres Seq. ID 1027153

<400>135  
 Phe Cys His Arg Ala Pro Leu Leu His Arg Thr Asp Pro Arg His  
 1 5 10 15  
 Phe Asp Tyr Thr Pro Ser Pro Ala Lys Met Val Thr Glu Lys Lys Thr  
 20 25 30  
 Lys Lys Ser His Glu Gly Ile Asn Ser Arg Leu Ala Leu Val Met Lys  
 35 40 45  
 Ser Gly Lys Tyr Thr Leu Gly Tyr Lys Ser Val Leu Lys Ser Leu Arg  
 50 55 60  
 Gly Ser Lys Gly Lys Leu Ile Leu Ile Ser Thr Asn Cys Pro Pro Leu  
 65 70 75 80  
 Arg Arg Ser Glu Ile Glu Tyr Tyr Ala Met Leu Ala Lys Val Gly Val  
 85 90 95  
 His His Tyr Asn Gly Asn Asn Val Asp Leu Gly Thr Ala Cys Gly Lys  
 100 105 110  
 Tyr Phe Arg Val Ser Cys Leu Ser Ile Val Asp Pro Gly Asp Ser Asp  
 115 120 125  
 Ile Ile Lys Ser Ile Pro Gly Asp Gln  
 130 135

<210> 136

<211> 112

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..112:Ceres Seq. ID 1027154

<400>136  
 Met Val Thr Glu Lys Lys Thr Lys Lys Ser His Glu Gly Ile Asn Ser  
 1 5 10 15  
 Arg Leu Ala Leu Val Met Lys Ser Gly Lys Tyr Thr Leu Gly Tyr Lys  
 20 25 30  
 Ser Val Leu Lys Ser Leu Arg Gly Ser Lys Gly Lys Leu Ile Leu Ile  
 35 40 45  
 Ser Thr Asn Cys Pro Pro Leu Arg Arg Ser Glu Ile Glu Tyr Tyr Ala  
 50 55 60  
 Met Leu Ala Lys Val Gly Val His Tyr Asn Gly Asn Asn Val Asp  
 65 70 75 80  
 Leu Gly Thr Ala Cys Gly Lys Tyr Phe Arg Val Ser Cys Leu Ser Ile  
 85 90 95  
 Val Asp Pro Gly Asp Ser Asp Ile Ile Lys Ser Ile Pro Gly Asp Gln  
 100 105 110

<210> 137

<211> 91

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..91:Ceres Seq. ID 1027155

<400>137  
 Met Lys Ser Gly Lys Tyr Thr Leu Gly Tyr Lys Ser Val Leu Lys Ser

1 Leu Arg Gly Ser Lys Gly Lys Leu Ile Leu Ile Ser Thr Asn Cys Pro 15  
 20 25 30  
 Pro Leu Arg Arg Ser Glu Ile Glu Tyr Tyr Ala Met Leu Ala Lys Val  
 35 40 45  
 Gly Val His Lys Tyr Asn Gly Asn Asn Val Asp Leu Gly Thr Ala Cys  
 50 55 60  
 Gly Lys Tyr Phe Arg Val Ser Cys Leu Ser Ile Val Asp Pro Gly Asp  
 65 70 75 80  
 Ser Asp Ile Ile Lys Ser Ile Pro Gly Asp Gln 90

<210> 138

<211> 653

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..653:Ceres Seq. ID 1028095

<400>138

aaataaagtt ttattattatc tatctatcca attggagacg aatcatatcg agcagtggaat 60  
 ttgcsagaac tggggcgcat ttccgacc ccgattcccg gcgcggtttt cgaacccggg 120  
 tgggtgtttt ggtcgagcgc cgtctgttgc agttccatcc agttccatcc cgtccattac 180  
 ctccccgga tattcgcttc ttccggagct ttgatgtcca attgctccag aanaagagac 240  
 attgattact ctcttcagca cgaagcgag tggagattga agctgtggtt ttcatagcg 300  
 tatgtctag cattgttgc cttaactgct tctgttggct tgcgtattca agattcggt 360  
 gtaagaactg ggccttcaac ttggactggt tgggctgggt tcttcaatg tgaatttga 420  
 ttgataagtg ggataatgta ttggacatcg cactacagat agcgacagc cacatactc 480  
 atctcgact cagattcata ctgtgacac ttcaagttc ctctcgttc tctactatgt 540  
 aaattccgt gtctgtgctt gtcaccttgg catgtgtgc gagattccic tgttgatca 600  
 gtgtgtgttt ctgttaaac ttgtgtgaa tgttattgaa aatactccaa atc 653

<210> 139

<211> 135

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..135:Ceres Seq. ID 1028096

<400>139

Met Asp Leu Ala Glu Leu Thr Ala Ile Phe Gly Pro Gly Phe Ser Gly  
 1 5 10 15  
 Ala Val Phe Gly Thr Gly Thr Phe Thr Val Asp Ala Val Val Cys  
 20 25 30  
 Ser Ser Ile Gln Val Pro Phe Val His Tyr Leu Pro Gly Ile Phe Ala  
 35 40 45  
 Ser Leu Gly Ala Leu Met Phe Asn Cys Val Arg Lys Glu Asp Ile Asp  
 50 55 60  
 Tyr Ser Pro Tyr Asp Glu Gly Glu Thr Arg Leu Lys Leu Thr Phe Phe  
 65 70 75 80  
 Ile Ala Tyr Val Val Ala Phe Val Ser Leu Ala Ala Ser Val Gly Leu  
 85 90 95  
 Leu Ile Gln Asp Ser Val Val Lys Thr Gly Pro Ser Thr Thr Thr Gly  
 100 105 110  
 Val Ala Gly Val Phe Gln Cys Val Phe Val Leu Ile Ser Gly Leu Met

115 120 125  
 Tyr Trp Thr Ser His Ser Glu  
 130 135  
 <210> 140  
 <211> 82  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..82:Ceres Seq. ID 1028097

<400>140

Met Phe Asn Cys Val Arg Lys Glu Asp Ile Asp Tyr Ser Pro Tyr Asp  
 1 5 10 15  
 Glu Gly Glu Trp Arg Leu Lys Leu Trp Leu Phe Ile Ala Tyr Val Val  
 20 25 30  
 Ala Phe Val Ser Leu Ala Ala Ser Val Gly Leu Leu Ile Gln Asp Ser  
 35 40 45  
 Val Val Lys Thr Gly Pro Ser Thr Trp Thr Gly Val Ala Gly Val Phe  
 50 55 60  
 Gln Cys Val Phe Val Leu Ile Ser Gly Leu Met Tyr Trp Thr Ser His  
 65 70 75 80  
 Ser Glu

<210> 141

<211> 80

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..80:Ceres Seq. ID 1028098

<400>141

Ile Lys Phe Leu Phe Ile Tyr Leu Ser Asn Trp Arg Arg Ile Ile Ser  
 1 5 10 15  
 Ser Asp Gly Phe Gly Arg Thr Val Gly Asp Phe Arg Thr Arg Ile Leu  
 20 25 30  
 Arg Arg Arg Phe Arg Asn Arg Val Val Leu Gly Arg Arg Arg  
 35 40 45  
 Leu Gln Phe His Pro Ser Ser Leu Arg Pro Leu Pro Ser Arg His Ile  
 50 55 60  
 Arg Phe Ser Arg Ser Phe Asp Val Gln Leu Arg Gln Lys Arg Arg His  
 65 70 75 80

<210> 142

<211> 779

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..779:Ceres Seq. ID 1028608

<400>142

gtctaaact taaactctaa agactttgtc attctgtgaa ccgtgtctcat aactgactctt 60  
 tccaaagat ctcaaaatc cctaaactc gaagtggca gaaccgcgc agacaatcaa 120  
 caccagatc tctatgcgc caccggaatc cgaagatccc accactattt ctgcaatgac 180  
 agatccaaca tctcaagaag cagctctcaa agacacagat ctgacgaag agcgcgaatc 240  
 agagaagaa ccaggaggaa tctctctcg aatctggcca ccgactcaga aaactcgca 300  
 ccgcttctg aatgcctga tgcagacct atccaccgaa tcaatctctt ctaagagata 360  
 ccgactctt aaactcgag atgcaaccac cgtccgnaa cttatcgaag aagagctta 420  
 tgggttgtc tgaatgctg tgcagtgga tgatgtggg attaagatc ttgagcttta 480  
 ttctaaagag attgaaaga ggaatgiga atctgttaag gtagatcta atgtatgtc 540  
 tggctctgt tccagagagg agaagtgga ggtctgagaa atttggtttt acttcttat 600  
 tataatctt tgcctctag ggtttgtgtg tcttgtaat gatgatgtg acttctgat 660  
 ttgtattat gcatattga agatgtttct tcttcttaa tatgagtga agtctcttc 720  
 779

&lt;210&gt; 143

&lt;211&gt; 180

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..180:Ceres Seq. ID 1028609

&lt;400&gt;143

Met Ala Glu Thr Ala Glu Thr Ile Asn Thr Thr Ile Ser Ser Pro Pro

1 5 10 15

Pro Glu Ser Glu Ser Ser Thr Thr Ile Ser Ala Met Thr Asp Pro Thr

20 25 30

Ser Glu Glu Ala Ala Ser Lys Asp Thr Asp Leu Thr Lys Glu Ala Glu

35 40 45

Ser Glu Lys Lys Pro Gly Gly Ile Ser Leu Arg Ile Trp Pro Pro Thr

50 55 60

Gln Lys Thr Arg Asp Ala Val Leu Asn Arg Leu Ile Glu Thr Leu Ser

65 70 75

Thr Glu Ser Ile Leu Ser Lys Arg Tyr Gly Thr Leu Lys Ser Asp Asp

85 90 95

Ala Thr Thr Val Ala Lys Leu Ile Glu Glu Ala Tyr Gly Val Ala

100 105 110

Ser Asn Ala Val Ser Ser Asp Asp Gly Ile Lys Ile Leu Glu Leu

115 120 125

Tyr Ser Lys Glu Ile Ser Lys Arg Met Leu Glu Ser Val Lys Ala Arg

130 135 140

Ser Asn Ala Ser Val Gly Asn Gly Ser Val Glu Asp Ala Asn Thr Asp

145 150 155

Ala Ser Glu Val Ser Lys Asp Ala Gly Pro Gly Ser Glu Glu Glu

165 170 175

Lys Ser Glu Ala

180

&lt;210&gt; 144

&lt;211&gt; 153

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..153:Ceres Seq. ID 1028610

&lt;400&gt;144

Met Thr Asp Pro Thr Ser Gln Glu Ala Ala Ser Lys Asp Thr Asp Leu 1  
 Thr Lys Glu Ala Glu Ser Glu Lys Lys Pro Gly Gly Ile Ser Leu Arg 15  
 20 25 30  
 Ile Trp Pro Pro Thr Gln Lys Thr Arg Asp Ala Val Leu Asn Arg Leu 35  
 40 45  
 Ile Glu Thr Leu Ser Thr Glu Ser Ile Leu Ser Lys Arg Tyr Gly Thr 50  
 55 60  
 Leu Lys Ser Asp Asp Ala Thr Thr Val Ala Lys Leu Ile Glu Glu Glu 65  
 70 75 80  
 Ala Tyr Gly Val Ala Ser Asn Ala Val Ser Ser Asp Asp Gly Ile 85  
 90 95  
 Lys Ile Leu Glu Leu Tyr Ser Lys Glu Ile Ser Lys Arg Met Leu Glu 100  
 105 110  
 Ser Val Lys Ala Arg Ser Asn Ala Ser Val Gly Asn Gly Ser Val Glu 115  
 120 125  
 Asp Ala Asn Thr Asp Ala Ser Glu Val Ser Lys Asp Ala Gly Pro 130  
 135 140  
 Gly Ser Glu Glu Lys Ser Glu Ala 145  
 150

&lt;210&gt; 145

&lt;211&gt; 94

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..94:Ceres Seq. ID 1028611

&lt;400&gt;145

Met Gln Pro Pro Ser Arg Asn Leu Ser Lys Lys Arg Leu Met Val Leu

1 5 10 15

Leu Arg Met Leu Cys Arg Val Met Met Gly Leu Arg Phe Leu Ser

20 25 30

Phe Ile Leu Lys Arg Leu Val Arg Gly Cys Leu Asn Leu Leu Arg Leu

35 40 45

Asp Leu Met Leu Val Leu Glu Met Glu Val Trp Arg Met Leu Ile Leu

50 55 60

Met Leu Val Arg Phe Leu Lys Met Met Leu Val Leu Val Gln Arg Arg

65 70 75 80

Arg Arg Val Arg Leu Glu Lys Phe Gly Phe Thr Ser Leu Leu

85 90

&lt;210&gt; 146

&lt;211&gt; 673

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..673:Ceres Seq. ID 1030069

&lt;400&gt;146

atccaaaaa aacaaaacaa aaaattatat tcaagagaaa aaggaaaaaa tgaattcat 60

ctccgatcag gtaagaagac tctcaagctc aacaccagag gagccagacc acaacagcc 120

agtcgaagga accgaagacag ctcaagacc agctaccac gccaggtcca tggcaatgc 180

caaggttgta gctgaagctg ctcaagccgc agctcgtaac gaacagaca aactcgaca 240

gggtaaagtc gccggagcct ctgctgatat cttagacgt tccgagaat acggttaagtt 300

cgatgaagaag agtagcaacty gtcaataact cgaagaagct gagaagtatc tcaacgaacta 360  
 cgaagtctca caactcacg gtgtgtgggg tectctect ccgacagagtc aggetdagcc 420  
 acaagtcag cctgagccgg cggctaagaa agacagatgaa gagtctgttg gtgggttgg 480  
 aggtatgcc aagatggctc aagttttctt gaagtatttt gacttttaatt tggttgtct 540  
 cattctcga ataataaatt aaataacag tatcgtttgt gactagtta tgtgtgtcg 600  
 ttatgttta tggggagtga cgaagtgtg taataacttc tggatgatcat gaactaatc 660  
 catctttgtt gtc 673

<210> 147

<211> 171

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..171:Ceres Seq. ID 1030070

<400>147

Ser Lys Lys Asn Lys Thr Lys Asn Tyr Ile Gln Glu Lys Lys Glu Lys 15

1 Met Asn Phe Ile Ser Asp Gln Val Lys Lys Leu Ser Ser Ser Thr Pro 20

Glu Glu Pro Asp His Asn Lys Pro Val Glu Gly Thr Glu Thr Ala Thr 25

35 Arg Pro Ala Thr Asn Ala Glu Leu Met Ala Ser Ala Lys Val Val Ala 40

50 Glu Ala Ala Gln Ala Ala Arg Asn Glu Ser Asp Lys Leu Asp Lys 55

65 Gly Lys Val Ala Gly Ala Ser Ala Asp Ile Leu Asp Ala Ser Glu Lys 60

85 Tyr Gly Lys Phe Asp Glu Lys Ser Ser Thr Gly His Tyr Leu Asp Lys 90

100 Ala Glu Lys Tyr Leu Asn Asp Tyr Glu Ser Ser His Ser Thr Gly Ala 105

115 Gly Gly Pro Pro Pro Thr Ser Gln Ala Glu Pro Ala Ser Gln Pro 120

130 Glu Pro Ala Ala Lys Lys Asp Asp Glu Ser Gly Gly Leu Lys 135

145 Gly Tyr Ala Lys Met Ala Gln Gly Phe Leu Lys 150

165 170

<210> 148

<211> 155

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..155:Ceres Seq. ID 1030071

<400>148

Met Asn Phe Ile Ser Asp Gln Val Lys Lys Leu Ser Ser Ser Thr Pro 10

1 Glu Glu Pro Asp His Asn Lys Pro Val Glu Gly Thr Glu Thr Ala Thr 15

20 Arg Pro Ala Thr Asn Ala Glu Leu Met Ala Ser Ala Lys Val Val Ala 25

35 Glu Ala Ala Gln Ala Ala Arg Asn Glu Ser Asp Lys Leu Asp Lys 40

55 60

Gly Lys Val Ala Gly Ala Ser Ala Asp Ile Leu Asp Ala Ser Glu Lys 70

65 Tyr Gly Lys Phe Asp Glu Lys Ser Ser Thr Gly His Tyr Leu Asp Lys 75

85 Ala Glu Lys Tyr Leu Asn Asp Tyr Glu Ser Ser His Ser Thr Gly Ala 90

100 Gly Gly Pro Pro Pro Thr Ser Gln Ala Glu Pro Ala Ser Gln Pro 105

115 Glu Pro Ala Ala Lys Lys Asp Asp Glu Glu Ser Gly Gly Leu Lys 120

130 Gly Tyr Ala Lys Met Ala Gln Gly Phe Leu Lys 135

145 155

<210> 149

<211> 115

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..115:Ceres Seq. ID 1030072

<400>149

Met Ala Ser Ala Lys Val Val Ala Glu Ala Ala Gln Ala Ala Ala Arg 10

1 Asn Glu Ser Asp Lys Leu Asp Lys Gly Lys Val Ala Gly Ala Ser Ala 20

25 Asp Ile Leu Asp Ala Ser Glu Lys Tyr Gly Lys Phe Asp Glu Lys Ser 30

35 Ser Thr Gly His Tyr Leu Asp Lys Ala Glu Lys Tyr Leu Asn Asp Tyr 40

50 Glu Ser Ser His Ser Thr Gly Ala Gly Gly Pro Pro Pro Thr Ser 55

65 Gln Ala Glu Pro Ala Ser Gln Pro Glu Pro Ala Ala Lys Lys Asp Asp 60

85 Glu Glu Ser Gly Gly Leu Gly Tyr Ala Lys Met Ala Gln Gly 90

100 Phe Leu Lys 105

115

<210> 150

<211> 944

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..944:Ceres Seq. ID 1032069

<400>150

atggcgaaatt ttaattgata actctctctc tctatctctg ttcaaacatc tctccaaag 60

taggtagaag aagatgggga cgagcttaga tgtatcaaga gcagagctag cactgtggt 120

aatgattttg aaccaagcag aggaagaga taagtattgc agagtatac agtatggttc 180

caagtcttgg agtgggtgac aacctgttac ttgtcaaat gttgacaaat ctactagctt 240

agcaagaaaa gtcttcgctc ttctcaagt ttgtgaatgac ttgatgctc ttactgctc 300

ttgtctaaa ggaactctc ttctctttg ttacttggg aagtcgaaga agcaactttt 360

atccaatc ttgtcttc atcaaatgt ctgctgtggg agatcaggaa tatataagaa 420

caagaacga gtgagttac ttgagctat atctcttc tgcgtgagg gatctctgt 480

ctgcacaact ttatcgagg ttgtagagt ggaaggctt tcttcatcaa tgaagaagt 540

cgaagaaggaa ctcaagaatg gaacaaagta tcaggatgag gattatcgtg ctaagctaaa 600  
 aaatcaaac gagagctaac ttcttttgaat caaatcagct atggacattg ttgagcagc 660  
 tggctctctt cagttagctc caacgaagt cactctcgt gtacccggag ctttggatt 720  
 catacctcc atcatttctt gttaccagtt gttccgaca cgcaccaaga ttaaaacacc 780  
 ctgaagtcaa tcaggaagc tgggttaag gagaagcca tttagaag ttctgtct 840  
 tctattatt aaggaaattc actgcgtgta taaaacatg tataaatac ttatttcag 900  
 gaattacgt tattataaat ttatgcctca gattctgat acgc 944

<210> 151  
 <211> 236  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..236:Ceres Seq. ID 1032070

<400>151  
 Met Gly Thr Thr Leu Asp Val Ser Arg Ala Glu Leu Ala Leu Val Val 15  
 1 5 10 15  
 Met Tyr Leu Asn Lys Ala Glu Ala Arg Asp Lys Leu Cys Arg Ala Ile 20  
 20 25 30  
 Gln Tyr Gly Ser Lys Phe Leu Ser Gly Gly Gln Pro Gly Thr Ala Gln 35  
 35 40 45  
 Asn Val Asp Lys Ser Thr Ser Leu Ala Arg Lys Val Phe Arg Leu Phe 50  
 50 55 60  
 Lys Phe Val Asn Asp Leu His Gly Leu Ile Ser Pro Val Pro Lys Gly 65  
 65 70 75 80  
 Thr Pro Leu Pro Leu Val Leu Leu Gly Lys Ser Lys Asn Ala Leu Leu 85  
 85 90 95  
 Ser Thr Phe Leu Phe Leu Asp Gln Ile Val Trp Leu Gly Arg Ser Gly 100  
 100 105 110  
 Ile Tyr Lys Asn Lys Glu Arg Ala Glu Leu Leu Gly Arg Ile Ser Leu 115  
 115 120 125  
 Phe Cys Trp Met Gly Ser Ser Val Cys Thr Thr Leu Val Glu Val Gly 130  
 130 135 140  
 Glu Met Gly Arg Leu Ser Ser Ser Met Lys Lys Ile Glu Lys Gly Leu 145  
 145 150 155  
 Lys Asn Gly Asn Lys Tyr Gln Asp Glu Asp Tyr Arg Ala Lys Leu Lys 160  
 160 165 170 175  
 Lys Ser Asn Glu Arg Ser Leu Ala Leu Ile Lys Ser Ala Met Asp Ile 180  
 180 185 190  
 Val Val Ala Ala Gly Leu Leu Gln Leu Ala Pro Thr Lys Ile Thr Pro 195  
 195 200 205  
 Arg Val Thr Gly Ala Phe Gly Phe Ile Thr Ser Ile Ile Ser Cys Tyr 210  
 210 215 220  
 Gln Leu Leu Pro Thr Arg Pro Lys Ile Lys Thr Pro 225  
 225 230 235

<210> 152  
 <211> 220  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..220:Ceres Seq. ID 1032071

<400>152

Met Tyr Leu Asn Lys Ala Glu Ala Arg Asp Lys Leu Cys Arg Ala Ile 5  
 5 10 15  
 Gln Tyr Gly Ser Lys Phe Leu Ser Gly Gly Gln Pro Gly Thr Ala Gln 20  
 20 25 30  
 Asn Val Asp Lys Ser Thr Ser Leu Ala Arg Lys Val Phe Arg Leu Phe 35  
 35 40 45  
 Lys Phe Val Asn Asp Leu His Gly Leu Ile Ser Pro Val Pro Lys Gly 50  
 50 55 60  
 Thr Pro Leu Pro Leu Val Leu Leu Gly Lys Ser Lys Asn Ala Leu Leu 65  
 65 70 75 80  
 Ser Thr Phe Leu Phe Leu Asp Gln Ile Val Trp Leu Gly Arg Ser Gly 85  
 85 90 95  
 Ile Tyr Lys Asn Lys Glu Arg Ala Glu Leu Leu Gly Arg Ile Ser Leu 100  
 100 105 110  
 Phe Cys Trp Met Gly Ser Ser Val Cys Thr Thr Leu Val Glu Val Gly 115  
 115 120 125  
 Glu Met Gly Arg Leu Ser Ser Met Lys Lys Ile Glu Lys Gly Leu 130  
 130 135 140  
 Lys Asn Gly Asn Lys Tyr Gln Asp Glu Asp Tyr Arg Ala Lys Leu Lys 145  
 145 150 155  
 Lys Ser Asn Glu Arg Ser Leu Ala Leu Ile Lys Ser Ala Met Asp Ile 165  
 165 170 175  
 Val Val Ala Ala Gly Leu Leu Gln Leu Ala Pro Thr Lys Ile Thr Pro 180  
 180 185 190  
 Arg Val Thr Gly Ala Phe Gly Phe Ile Thr Ser Ile Ile Ser Cys Tyr 195  
 195 200 205  
 Gln Leu Leu Pro Thr Arg Pro Lys Ile Lys Thr Pro 210  
 210 215 220

<210> 153

<211> 769

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..769:Ceres Seq. ID 1033557

<400>153  
 ggtatcgtcc gtagacgata gaggtccac actatgtaaa ggagaccgga gaacaaagg 60  
 aactgggttt ggagataaga agtgggaact tgatggctt cctatccac ttccgttga 120  
 gcttcggtt cgtcgcgttt atggatcct gctgcctcca acgcaagat ttgcgttct 180  
 tctgtctgc ttctctgcg tacgggttg agcggagtt cgtctctct caattctcc 240  
 gctcttcgc aattgtcca ttctcgttt ctctcgtgc ctgttccct agcgttcca 300  
 ttcttcggt tgcctatgc gttgatctc agcagtcaca ctatggact gaatggcag 360  
 agacgagag gcttctggt tagactgga aagctgctc tgtgtcaaac taagagaagc 420  
 agatcaaea aatcttagc taggactcat gtttcgca gaagatgag gaccactagc 480  
 ggtagacaa ccataaagc tcgaagtcg aaggagcgt ggaacctcg tcccaagtc 540  
 aacctagca gggcaaacg ggtctgaagt ttgttctct ctcaactca tctgaactc 600  
 tgtttcatg gcttttttc caccagaaga tacaatgaga aacatggtc ttgagctct 660  
 ttctcgtt ttaactctc tgatccttg attcgttat aaatcaatg tcttgaatt 720  
 catacatgt actactatac tacttgcttc attgaattg ttactact 769

<210> 154

<211> 157

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..157:Ceres Seq. ID 1033558

&lt;400&gt;154

Met Ala Ser Leu Ser Thr Ser Val Val Ala Ser Ala Ser Ser Arg Leu  
 1 5 10 15  
 Trp Asn Pro Ala Ala Ser Asn Gly Lys Ile Cys Val Pro Ser Ala Ser  
 20 25 30  
 Leu Ser Leu Arg Thr Gly Cys Arg Ser Ser Ser Ser Leu Thr Ser  
 35 40 45  
 Ser Ala Ser Ser Gln Leu Leu His Cys Ser Phe Leu Ser Ser Pro Val  
 50 55 60  
 Ser Leu Ala Ser Pro Phe Ser Gly Leu Ser Ile Ala Phe Asp Leu Ser  
 65 70 75  
 Ser Gln Thr Ser Gly Leu Asn Gly Gln Arg Arg Arg Gly Leu Val Val  
 80 85 90 95  
 Arg Ala Gly Lys Ala Ala Leu Cys Gln Thr Lys Arg Ser Arg Ser Arg  
 100 110  
 Lys Ser Leu Ala Arg Thr His Gly Phe Arg Arg Arg Met Arg Thr Thr  
 115 120 125  
 Ser Gly Arg Ala Thr Ile Lys Arg Arg Ala Lys Gly Arg Trp Asn  
 130 135 140  
 Leu Cys Pro Lys Ser Asn Pro Ser Ser Gly Lys Arg Ala  
 145 150 155

&lt;210&gt; 155

&lt;211&gt; 679

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..679:Ceres Seq. ID 1034688

&lt;400&gt;155

aaaaaaac aaacaaaaa attatattca agagaaaag gaataaatga atttcatctc  
 60  
 cgaatcggta agaaactct caagctaac accnagggag ccagaccaca acaagccagt  
 120  
 cgaagaaac gaacacagta caagccagc taccaacgc gactcatgg caagtgcaca  
 180  
 gggttagct gaagctgctc aagccgagc tggtaecgaa tcagacaac tcgacaagg  
 240  
 taagtcgcc ggagccttg ctgatatctt agagctgccc gagaatacg gtaattcga  
 300  
 tgaagaagt agcactggtc agtacctcga caagctgag agtatactca acgactcga  
 360  
 gtcgtcacac tccaccggtg ctgggtgtcc tctctcca acgagtcagg ctgagccagc  
 420  
 aagtcagcct gagccgagg ctaagaaga cgaatgaag tctgggttg ggtctgagg  
 480  
 ttctgtaata ataaattaaa taactagtat cgtttgtgac tagttatgt tgcctcgtt  
 540  
 atgttatgg ggagtgacga gtgagtgtaa taactcttgg tgaatcaga tctatccat  
 600  
 cttgtgttg attatgctc  
 660  
 679

&lt;210&gt; 156

&lt;211&gt; 170

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..170:Ceres Seq. ID 1034689

&lt;400&gt;156

Lys Lys Asn Lys Thr Lys Asn Tyr Ile Gln Glu Lys Lys Glu Lys Met

1 5 10 15  
 Asn Phe Ile Ser Asp Gln Val Lys Lys Leu Ser Ser Thr Pro Glu  
 20 25 30  
 Glu Pro Asp His Asn Lys Pro Val Glu Gly Thr Glu Thr Ala Thr Arg  
 35 40 45  
 Pro Ala Thr Asn Ala Glu Leu Met Ala Ser Ala Lys Val Val Ala Glu  
 50 55 60  
 Ala Ala Gln Ala Ala Arg Asn Glu Ser Asp Lys Leu Asp Lys Gly  
 65 70 75  
 Lys Val Ala Gly Ala Ser Ala Asp Ile Leu Asp Ala Ala Glu Lys Tyr  
 80 85 90 95  
 Gly Lys Phe Asp Glu Lys Ser Ser Thr Gly Gln Tyr Leu Asp Lys Ala  
 100 105 110  
 Glu Lys Tyr Leu Asn Asp Tyr Glu Ser Ser His Ser Thr Gly Ala Gly  
 115 120 125  
 Gly Pro Pro Pro Thr Ser Gln Ala Glu Pro Ala Ser Gln Pro Glu  
 130 135 140  
 Pro Ala Ala Lys Lys Asp Asp Glu Ser Gly Gly Leu Gly Gly  
 145 150 155 160  
 Tyr Ala Lys Met Ala Gln Gly Phe Leu Lys  
 165 170

&lt;210&gt; 157

&lt;211&gt; 155

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..155:Ceres Seq. ID 1034690

<400>157  
 1 5 10 15  
 Met Asn Phe Ile Ser Asp Gln Val Lys Lys Leu Ser Ser Thr Pro  
 1 5 10 15  
 Glu Glu Pro Asp His Asn Lys Pro Val Glu Gly Thr Glu Thr Ala Thr  
 20 25 30  
 Arg Pro Ala Thr Asn Ala Glu Leu Met Ala Ser Ala Lys Val Val Ala  
 35 40 45  
 Glu Ala Ala Gln Ala Ala Arg Asn Glu Ser Asp Lys Leu Asp Lys  
 50 55 60  
 Gly Lys Val Ala Gly Ala Ser Ala Asp Ile Leu Asp Ala Ala Glu Lys  
 65 70 75  
 Tyr Gly Lys Phe Asp Glu Lys Ser Ser Thr Gly Gln Tyr Leu Asp Lys  
 80 85 90 95  
 Ala Glu Lys Tyr Leu Asn Asp Tyr Glu Ser Ser His Ser Thr Gly Ala  
 100 105 110 115  
 Gly Gly Pro Pro Thr Ser Gln Ala Glu Pro Ala Ser Gln Pro  
 120 125  
 Glu Pro Ala Ala Lys Lys Asp Asp Glu Ser Gly Gly Gly Leu Gly  
 130 135 140  
 Gly Tyr Ala Lys Met Ala Gln Gly Phe Leu Lys  
 145 150 155

&lt;210&gt; 158

&lt;211&gt; 115

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..115:Ceres Seq. ID 1034691

<400>158  
Met Ala Ser Ala Lys Val Val Ala Glu Ala Ala Gln Ala Ala Arg  
1 5 10 15  
Asn Glu Ser Asp Lys Leu Asp Lys Gly Lys Val Ala Gly Ala Ser Ala  
20 25 30  
Asp Ile Leu Asp Ala Ala Glu Lys Tyr Gly Lys Phe Asp Glu Lys Ser  
35 40 45  
Ser Thr Gly Gln Tyr Leu Asp Lys Ala Glu Lys Tyr Leu Asn Asp Tyr  
50 55 60  
Glu Ser Ser His Ser Thr Gly Ala Gly Gly Pro Pro Pro Thr Ser  
65 70 75  
Gln Ala Glu Pro Ala Ser Gln Pro Glu Pro Ala Ala Lys Lys Asp Asp  
85 90 95  
Glu Glu Ser Gly Gly Gly Leu Gly Tyr Ala Lys Met Ala Gln Gly  
100 105 110  
Phe Leu Lys  
115

&lt;210&gt; 159

&lt;211&gt; 471

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..471:Ceres Seq. ID 1035033

<400>159  
ataaataaaa aaaacttatt ttgtctatc tttaagttt aaasacagtc aattcattta  
tatcatctat gatgagaaaa ctcatcaag tctgtctcac tgttatgatt atcttcacca  
ttctgtgct agggagggg gcaaatgagg ggcctaggaaa accaaaaaaa caatgtaatg  
agattctaaa gcaatccaac tgggttgctg cagatgtgga cttctatggt gtgagaaga  
ggggaaagg agccgggttac tgtctcctt ctaaaagtg ttactgctat tatcatgcc  
cttaacgatt ttattattta caataataag agattataga taaagtatga aaataatgaa  
tcgggtggtt atagcccaac gattgataa gttgtctcta tcaataaac tatatagact  
aaaagtaata ataataatga caatcataac tcagaagttt ttgttatgic c  
471

&lt;210&gt; 160

&lt;211&gt; 100

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..100:Ceres Seq. ID 1035034

<400>160  
Ile Lys Lys Lys Lys Leu Ile Leu Phe Ile Leu Leu Ser Leu Lys Thr Val  
1 5 10 15  
Asn Ser Phe Ile Ser Ser Met Met Lys Lys Leu Ile Gln Val Ser Phe  
20 25 30  
Thr Val Met Ile Ile Phe Thr Ile Leu Val Leu Gly Val Val Ala Asn  
35 40 45  
Glu Gly Leu Gly Lys Pro Lys Lys Gln Cys Asn Glu Ile Leu Lys Gln  
50 55 60  
Ser Asn Cys Val Ala Ala Glu Cys Asp Ser Met Cys Val Lys Lys Arg

65 Gly Lys Gly Ala Gly Tyr Cys Ser Pro Ser Lys Lys Cys Tyr Cys Tyr  
70 75 80  
Tyr His Cys Pro  
85 90 95 100

&lt;210&gt; 161

&lt;211&gt; 78

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..78:Ceres Seq. ID 1035035

<400>161  
Met Met Lys Lys Lys Leu Ile Gln Val Ser Phe Thr Val Met Ile Ile Phe  
1 5 10 15  
Thr Ile Leu Val Leu Gly Val Val Ala Asn Glu Gly Leu Gly Lys Pro  
20 25 30  
Lys Lys Gln Cys Asn Glu Ile Leu Lys Gln Ser Asn Cys Val Ala Ala  
35 40 45  
Glu Cys Asp Ser Met Cys Val Lys Lys Arg Gly Lys Gly Ala Gly Tyr  
50 55 60  
Cys Ser Pro Ser Lys Lys Cys Tyr Cys Tyr Tyr His Cys Pro  
65 70 75

&lt;210&gt; 162

&lt;211&gt; 77

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..77:Ceres Seq. ID 1035036

<400>162  
Met Lys Lys Lys Leu Ile Gln Val Ser Phe Thr Val Met Ile Ile Phe Thr  
1 5 10 15  
Ile Leu Val Leu Gly Val Val Ala Asn Glu Gly Leu Gly Lys Pro Lys  
20 25 30  
Lys Gln Cys Asn Glu Ile Leu Lys Gln Ser Asn Cys Val Ala Ala Glu  
35 40 45  
Cys Asp Ser Met Cys Val Lys Lys Arg Gly Lys Gly Ala Gly Tyr Cys  
50 55 60  
Ser Pro Ser Lys Lys Cys Tyr Cys Tyr Tyr Tyr His Cys Pro  
65 70 75

&lt;210&gt; 163

&lt;211&gt; 985

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..985:Ceres Seq. ID 1035071

&lt;400&gt;163



taacgaactc tcttctccga atctagggtt ttgtctctca gctcctccaa gctttgcaaa 60  
 caatcaggc gaccagatc agacagaaga gaaggttcgt agccgatggt gtgttttaccg 120  
 ctgaactcaa tggagttcta accagagagc ttgtgagga tggttactct ggtgttgagg 180  
 ttgcgtcac tccatgctgt actgagatca tctatagagc cactcgtact caaacggttc 240  
 tccgtgagaa gggagggggg attagggaat tgaactctct tgcacaaag agattcfaat 300  
 ttctcagga caggtttgag ctttatgctg agaggtgttc taacagaggt ctttgtgcta 360  
 ttgtcaggc tgaatctctt cgttcaagc tcttgggtg tcttggctgt cgtaggcgct 420  
 gctatggtt tttaaggtt gttatggaga gtggagctaa aggatggaaa gtcattgtga 480  
 gtgaaaaat ccgtgtgca cgtgtgaatg agatgaagtt caaggtatgt tacatggttt 540  
 ctcagatca gcaaccaag gaatacatg atgtgcagtt taggcattgt cttcttagac 600  
 aggtgtgct tgtgtccaa gtgaagatca tctgtgactg ggaccacaag ggcacaacag 660  
 gaccgtgac accattgctt gttgttga tctctctatc accgaagaa gatgatgtat 720  
 acattgacc tgcctagtt gttatcaag ctgtcttctg agagagcctt ttcgtttctt ccatatacca 780  
 ccaagatca tccagcaatg ccagttgctt agagagcctt ttcgtttctt ggtttaaaac 840  
 ccaatggga tggttcttt ttgggacag tactagtttt atcatgttat tctaagttaa 900  
 tacttaagt tctggaggtt gaattacatc aaggatatga atgtatggtt ttggttatct 960  
 ttgtttaaa tttttttat ctatt 985

&lt;210&gt; 164

&lt;211&gt; 148

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..148:Ceres Seq. ID 1035072

&lt;400&gt;164

Met Ala Thr Gln Ile Ser Lys Lys Arg Lys Phe Val Ala Asp Gly Val

1 5 10 15

Phe Tyr Ala Glu Leu Asn Glu Val Leu Thr Arg Glu Leu Ala Glu Asp

20 25 30

Gly Tyr Ser Gly Val Glu Val Arg Val Thr Pro Met Arg Thr Glu Ile

35 40 45

Ile Ile Arg Ala Thr Arg Thr Gln Asn Val Leu Gly Glu Lys Gly Arg

50 55 60

Arg Ile Arg Glu Leu Thr Ser Leu Val Gln Lys Arg Phe Lys Phe Pro

65 70 75 80

Gln Asp Ser Val Glu Leu Tyr Ala Glu Lys Val Ala Asn Arg Gly Leu

85 90 95

Cys Ala Ile Ala Gln Ala Glu Ser Leu Arg Tyr Lys Leu Leu Gly Gly

100 105 110

Leu Ala Val Arg Arg Ala Cys Tyr Gly Val Leu Arg Phe Val Met Glu

115 120 125

Ser Gly Ala Lys Gly Cys Glu Val Ile Val Ser Gly Lys Leu Arg Ala

130 135 140

Ala Arg Ala Lys

145

&lt;210&gt; 165

&lt;211&gt; 105

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..105:Ceres Seq. ID 1035073

&lt;400&gt;165

Met Arg Thr Glu Ile Ile Arg Ala Thr Arg Thr Gln Asn Val Leu 15  
 1 5 10  
 Gly Glu Lys Gly Arg Arg Ile Arg Glu Leu Thr Ser Leu Val Gln Lys 20  
 25 30  
 Arg Phe Lys Phe Pro Gln Asp Ser Val Glu Leu Tyr Ala Glu Lys Val 35  
 40 45  
 Ala Asn Arg Gly Leu Cys Ala Ile Ala Gln Ala Glu Ser Leu Arg Tyr 50  
 55 60  
 Lys Leu Leu Gly Gly Leu Ala Val Arg Arg Ala Cys Tyr Gly Val Leu 65  
 70 75 80  
 Arg Phe Val Met Glu Ser Gly Ala Lys Gly Cys Glu Val Ile Val Ser 95  
 100

&lt;210&gt; 166

&lt;211&gt; 99

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..99:Ceres Seq. ID 1035074

&lt;400&gt;166

Met Lys Phe Lys Asp Gly Tyr Tyr Met Val Ser Ser Gly Gln Pro Thr Lys

1 5 10 15

Glu Tyr Ile Asp Ala Ala Val Arg His Val Leu Leu Arg Gln Gly Val

20 25 30

Leu Gly Leu Lys Val Lys Ile Met Leu Asp Trp Asp Pro Lys Gly Lys

35 40 45

Gln Gly Pro Met Thr Pro Leu Pro Asp Val Val Ile Ile His Thr Pro

50 55 60

Lys Glu Asp Asp Val Tyr Ile Ala Pro Ala Gln Val Val Thr Gln Ala

65 70 75 80

Ala Phe Val Pro Glu Ala Pro Leu Thr Thr Thr Asp Tyr Pro Ala Met

85 90 95

Pro Val Ala

&lt;210&gt; 167

&lt;211&gt; 759

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..759:Ceres Seq. ID 1376587

&lt;400&gt;167

atcccgagaa aaccctattg ctccctttg ctcaagctca gtgcctcttt ggaagcaaa 60

ctcaaccaac aagacgatga ttatctaga gaacaatcgc agagagatct gcaaatacct 120

tttcaagag ggaatagct ttgctaagaa ggatttcaat ctccaagac atccgtgat 180

tgatgacca aacctgcaag tgattaagct catcgagagt ttcaagtcca aggaatcgt 240

tagggagaca ttgctggga tgcattatta ttggtttctg actaatgaag gaattgagt 300

cttgagaact tatcttaacc ttcctccga tgtgtccct gctacttga agaactctgc 360

taagcccggt ggtgcctcct ttggtggccc acctggtgat cgccaaagag gaccacctcg 420

ctctgagga gaccgtcca gatttgtga taccgtatga taccgtatga gcccacgag 480

tggtgatgag aagagtgag ctccagctga ttccagccg tcttccaag gaggtggtgg 540

tagcctgtg ttggccgtg gtccagcgg ttacagtga gcagccat ctggtcagg 600  
gtcccttga aaatttgg ttcatatgc gacatggaa gcacagttt gtttttgg 660  
ctagttgt ttgttaag caaatcga actataac tatctattac ctccactg 720  
gtttaaacga acaaatcctg tagttcagg attagtgc 759

<210> 168

<211> 202

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..202:Ceres Seq. ID 1376588

<400>168

Ser Ala Glu Asn Pro Ile Ala Pro Leu Cys Ser Ser Val Pro Leu  
1 5 10 15

Cys Ser Glu Ser Ser Thr Asn Lys Thr Met Ile Ile Ser Glu Asn Asn  
20 25 30

Arg Arg Glu Ile Cys Lys Tyr Leu Phe Lys Glu Gly Val Cys Phe Ala  
35 40 45

Lys Lys Asp Phe Asn Leu Pro Lys His Pro Leu Ile Asp Val Pro Asn  
50 55 60

Leu Gln Val Ile Lys Leu Met Gln Ser Phe Lys Ser Lys Glu Tyr Val  
65 70 75 80

Arg Glu Thr Phe Ala Trp Met His Tyr Tyr Trp Phe Leu Thr Asn Glu  
85 90 95

Gly Ile Glu Phe Leu Arg Thr Tyr Leu Asn Leu Pro Ser Asp Val Val  
100 105 110

Pro Ala Thr Leu Lys Lys Ser Ala Lys Pro Gly Gly Arg Pro Phe Gly  
115 120 125

Gly Pro Pro Gly Asp Arg Gln Arg Gly Pro Pro Arg Ser Asp Gly Asp  
130 135 140

Arg Pro Arg Phe Gly Asp Arg Asp Gly Tyr Arg Gly Gly Pro Arg Gly  
145 150 155 160

Gly Asp Glu Lys Ser Gly Ala Pro Ala Asp Phe Gln Pro Ser Phe Gln  
165 170 175

Gly Gly Gly Arg Pro Gly Phe Gly Arg Gly Ala Gly Gly Tyr Ser  
180 185 190

Ala Ala Ala Pro Ser Gly Ser Gly Phe Pro  
195 200

<210> 169

<211> 177

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..177:Ceres Seq. ID 1376589

<400>169

Met Ile Ile Ser Glu Asn Asn Arg Arg Glu Ile Cys Lys Tyr Leu Phe  
1 5 10 15

Lys Glu Gly Val Cys Phe Ala Lys Lys Asp Phe Asn Leu Pro Lys His  
20 25 30

Pro Leu Ile Asp Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln Ser  
35 40 45

Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His Tyr

Tyr Trp Phe Leu Thr Asn Glu Gly Ile Glu Phe Leu Arg Thr Tyr Leu  
50 55 60  
65 70 75  
Asn Leu Pro Ser Asp Val Val Pro Ala Thr Leu Lys Lys Ser Ala Lys  
85 90 95

Pro Gly Gly Arg Pro Phe Gly Gly Pro Pro Gly Asp Arg Gln Arg Gly  
100 105 110

Pro Pro Arg Ser Asp Gly Asp Arg Pro Arg Phe Gly Asp Arg Asp Gly  
115 120 125

Tyr Arg Gly Gly Pro Arg Gly Gly Asp Glu Lys Ser Gly Ala Pro Ala  
130 135 140

Asp Phe Gln Pro Ser Phe Gln Gly Gly Arg Pro Gly Phe Gly  
145 150 155

Arg Gly Ala Gly Tyr Ser Ala Ala Pro Ser Gly Ser Gly Phe  
165 170 175

Pro

<210> 170

<211> 132

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..132:Ceres Seq. ID 1376590

<400>170

Met Gln Ser Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp  
1 5 10 15

Met His Tyr Tyr Trp Phe Leu Thr Asn Glu Gly Ile Glu Phe Leu Arg  
20 25 30

Thr Tyr Leu Asn Leu Pro Ser Asp Val Val Pro Ala Thr Leu Lys Lys  
35 40 45

Ser Ala Lys Pro Gly Gly Arg Pro Phe Gly Gly Pro Pro Gly Asp Arg  
50 55 60

Gln Arg Gly Pro Pro Arg Ser Asp Gly Asp Arg Pro Arg Phe Gly Asp  
65 70 75 80

Arg Asp Gly Tyr Arg Gly Gly Pro Arg Gly Asp Glu Lys Ser Gly  
85 90 95

Ala Pro Ala Asp Phe Gln Pro Ser Phe Gln Gly Gly Gly Arg Pro  
100 105 110

Gly Phe Gly Arg Gly Ala Gly Tyr Ser Ala Ala Pro Ser Gly  
115 120 125

Ser Gly Phe Pro

130

<210> 171

<211> 645

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..645:Ceres Seq. ID 1378581

<400>171

gattcccat ttgtgttga gcgatcgag agctctgacg gcaccagcca cctctgcgc 60  
ttgcgaatc caataagct gtggatgaaa cgaagaatcc tagatttcg aatggcgaca 120

ccatttatag cgggggttgc agtagctgca acagcacttg ctggtagata tggatccaa 180  
 gcatggcaag catcgaagc agggccacca aggcccaaaa ttaagaatt ttatggggc 240  
 gtttcagc ctagcatgac gaagaggaa gctgctctca ttcttgctt ctagggagat 300  
 gtacggcgg aagagtgaa ggaagcac agbaagtaa tggtagcaaa caatccagat 360  
 gcagggtga gccatttct agctcttaag atcaatgaag cttaagacgt gatgctaggc 420  
 aaaaactaaa acaggggac cgtcttttga tgcnaattc aatcaaaag cttccttatg 480  
 aaactactta tgnataggc ctgtaatgac aaatgcaga tttaattga attataagc 540  
 ttgtgtctg ttgtgacac ataaagactc gaactcttc ttcaacttgt tgcgtgtaa 600  
 aagaacacaa aaataaaac gcaagtcaga agggcgttct agtcg 645

&lt;210&gt; 172

&lt;211&gt; 112

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..112:Ceres Seq. ID 1378582

&lt;400&gt;172

Met Ala Thr Pro Phe Ile Ala Gly Val Ala Val Ala Ala Thr Ala Leu 1  
 Ala Gly Arg Tyr Gly Ile Gln Ala Trp Gln Ala Phe Lys Ala Arg Pro 20  
 Pro Arg Pro Lys Ile Lys Lys Phe Tyr Gln Gly Gly Phe Gln Pro Thr 35  
 Met Thr Lys Arg Glu Ala Ala Leu Ile Leu Gly Val Arg Glu Ser Val 50  
 Ala Ala Glu Lys Val Lys Glu Ala His Arg Lys Val Met Val Ala Asn 65  
 His Pro Asp Ala Gly Gly Ser His Phe Leu Ala Ser Lys Ile Asn Glu 85  
 Ala Lys Asp Val Met Leu Gly Lys Thr Lys Asn Ser Gly Ser Ala Phe 100

&lt;210&gt; 173

&lt;211&gt; 64

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..64:Ceres Seq. ID 1378583

&lt;400&gt;173

Met Thr Lys Arg Glu Ala Ala Leu Ile Leu Gly Val Arg Glu Ser Val 1  
 Ala Ala Glu Lys Val Lys Glu Ala His Arg Lys Val Met Val Ala Asn 20  
 His Pro Asp Ala Gly Gly Ser His Phe Leu Ala Ser Lys Ile Asn Glu 35  
 Ala Lys Asp Val Met Leu Gly Lys Thr Lys Asn Ser Gly Ser Ala Phe 50

&lt;210&gt; 174

&lt;211&gt; 803

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..803:Ceres Seq. ID 1383462

&lt;400&gt;174

aggaacagtt tctatctaga gagaacacaa cacatcgaaac catggagaaa tcaacaagaa 60  
 cccattatcat aaccatgtg ataactcca tgttgctagg gtttgaaac tctgatctt 120  
 ctcagagcag agagagagtt acgaaccagc ctagagact atccacgtgt atccgtact 180  
 ttgagagaga cgcaagagtt ccaacaaaag attgtgtgc aggtttggc caagttaa 240  
 gaagagtga gaegtgtt tgcatttgg tcsagacaaa agatgatct caacttgta 300  
 tcaagattaa cgaacccta gccctcac ttccctcgc ttgtcatatc agcgtccta 360  
 acatcactga ctgtatttgg attctgcatc tactctgtaa ctcnacattg gctaaagat 420  
 ttgagaactt aggaaggatt gaagacaact acaactcac atctctaca caattcata 480  
 aagatgggac aggagaggga aagccgaac cagtgaagag taatggatcg aaggaga 540  
 gtgtgtggg tgttgagctt ttaataatt tgcattgttc tctcatcttc tttaattac 600  
 ttccttctt taatgttatt ttgtttatt attacaataa cttataagat atctattat 660  
 atatgagaca catgcagtgt ttgcgatct atatgatgt aaatagtc aatgtagta 720  
 ttaatacag aagcgtacat ttaatttgg ttgtgcaa ttgatctgt acgttagtga 780  
 atttatgga ttatcgactt act 803

&lt;210&gt; 175

&lt;211&gt; 197

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..197:Ceres Seq. ID 1383463

&lt;400&gt;175

Asp Thr Phe Leu Ser Arg Glu Asn Thr Thr His Arg Thr Met Glu Lys 1  
 Ser Thr Arg Thr Leu Phe Ile Thr Ile Val Ile Thr Ser Met Leu Leu 20  
 Gly Phe Gly Asn Ser Asp Leu Ala Gln Asp Arg Glu Glu Cys Thr Asn 35  
 Gln Leu Ile Glu Leu Ser Thr Cys Ile Pro Tyr Val Gly Gly Asp Ala 50  
 Lys Ala Pro Thr Lys Asp Cys Cys Ala Gly Phe Gly Gln Val Ile Arg 65  
 Lys Ser Glu Lys Cys Val Cys Ile Leu Val Arg Asp Lys Asp Asp Pro 85  
 Gln Leu Gly Ile Lys Ile Asn Ala Thr Leu Ala Ala His Leu Pro Ser 100  
 Ala Cys His Ile Thr Ala Pro Asn Ile Thr Asp Cys Ile Ser Ile Leu 110  
 His Leu Pro Arg Asn Ser Thr Leu Ala Lys Glu Phe Glu Asn Leu Gly 130  
 Arg Ile Glu Asp Asn Tyr Asn Ser Thr Ser Pro Thr Gln Ile His Lys 140  
 145 Asp Gly Thr Gly Gly Lys Ala Glu Pro Val Lys Ser Asn Gly Trp 155  
 Lys Glu Lys Ser Trp Leu Gly Val Glu Leu Leu Ile Tyr Leu Leu Val 175  
 Ser Leu Ile Phe Phe 185  
 195  
 <210> 176  
 <211> 184

<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..184:Ceres Seq. ID 1383464

<400>176

Met Glu Lys Ser Thr Arg Thr Leu Phe Ile Thr Ile Val Ile Thr Ser  
1 5 10 15  
Met Leu Leu Gly Phe Gly Asn Ser Asp Leu Ala Gln Asp Arg Glu Glu  
20 25 30  
Cys Thr Asn Gln Leu Ile Glu Leu Ser Thr Cys Ile Pro Tyr Val Gly  
35 40 45  
Gly Asp Ala Lys Ala Pro Thr Lys Asp Cys Cys Ala Gly Phe Gly Gln  
50 55 60  
Val Ile Arg Lys Ser Glu Lys Cys Val Cys Ile Leu Val Arg Asp Lys  
65 70 75 80  
Asp Asp Pro Gln Leu Gly Ile Lys Ile Asn Ala Thr Leu Ala Ala His  
85 90 95  
Leu Pro Ser Ala Cys His Ile Thr Ala Pro Asn Ile Thr Asp Cys Ile  
100 105 110  
Ser Ile Leu His Leu Pro Arg Asn Ser Thr Leu Ala Lys Glu Phe Glu  
115 120 125  
Asn Leu Gly Arg Ile Glu Asp Asn Tyr Asn Ser Thr Ser Pro Thr Gln  
130 135 140  
Ile His Lys Asp Gly Thr Gly Gly Lys Ala Glu Pro Val Lys Ser  
145 150 155  
Asn Gly Trp Lys Glu Lys Ser Trp Leu Gly Val Glu Leu Ile Tyr  
165 170 175  
Leu Leu Val Ser Leu Ile Phe Phe  
180

<210> 177

<211> 168

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..168:Ceres Seq. ID 1383465

<400>177

Met Leu Leu Gly Phe Gly Asn Ser Asp Leu Ala Gln Asp Arg Glu Glu  
1 5 10 15  
Cys Thr Asn Gln Leu Ile Glu Leu Ser Thr Cys Ile Pro Tyr Val Gly  
20 25 30  
Gly Asp Ala Lys Ala Pro Thr Lys Asp Cys Cys Ala Gly Phe Gly Gln  
35 40 45  
Val Ile Arg Lys Ser Glu Lys Cys Val Cys Ile Leu Val Arg Asp Lys  
50 55 60  
Asp Asp Pro Gln Leu Gly Ile Lys Ile Asn Ala Thr Leu Ala Ala His  
65 70 75 80  
Leu Pro Ser Ala Cys His Ile Thr Ala Pro Asn Ile Thr Asp Cys Ile  
85 90 95  
Ser Ile Leu His Leu Pro Arg Asn Ser Thr Leu Ala Lys Glu Phe Glu  
100 105 110  
Asn Leu Gly Arg Ile Glu Asp Asn Tyr Asn Ser Thr Ser Pro Thr Gln  
115 120 125

Ile His Lys Asp Gly Thr Gly Gly Lys Ala Glu Pro Val Lys Ser  
130 135 140  
Asn Gly Trp Lys Glu Lys Ser Trp Leu Gly Val Glu Leu Ile Tyr  
145 150 155  
Leu Leu Val Ser Leu Ile Phe Phe  
165

<210> 178

<211> 968

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..968:Ceres Seq. ID 1386215

<400>178

aagagaaaa caaaactgca aaacaaaaa taaaaaaac atcgacaag aaaataaaag  
attgtagaa tcaactaaga aaatgctag cactatgat actacattgc ctcaagtcaa  
tggcttcga gcaacaaaa tctctgcgc tctgtacaa ggcctggcaa gtgtcagcc  
catgagcgc aagggaatg gagcttggg tgcagaatg gacttcacg gtccatcac  
aaatctgata atgttaacgt cagcacccg gatgtgttc gcggggagat tcgactatgc  
gccatcaccc ataggaaagg ccacagctgg acttagttg gaggcacgtg actcagttc  
aacaacggt gaccggcgg ggttcacgt tgcgacact ttgcttgg cctccaatt  
aacccaatt tcaacttct cccaattc cgatcacgt agacgtccc cgaatgaacg  
ctctcactc cgaatcggc aatggctctt gtccagctg gaaatattc ggtggcta  
atgaataag cgtgtgctct gttaaaaggaa gaagctctta agaagatcat agcagttaa  
acggcaagt tgaagctaa aggtgcattg acgaatttc aggtattgt tattcgacg  
agtaaggcca cggcaagag gaagaaagt gaacgtgat ttctaaatc ttatagcaa  
tttactggt acgtagagct ttgtgtgat atacgtctc taggttcaa ggaagaacta  
cttttttgc gttatggcc attatctt gtgaattta gtccattta ggcgttttat  
gatcaatgt ctccagaga aagctggaga cttgaatta ggataatta ggcgttttat  
ttgtacaat ggtgtgtga agactcagt acattcattc atattaatg gtccattt  
caatcttt

<210> 179

<211> 180

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..180:Ceres Seq. ID 1386216

<400>179

Arg Gly Lys Gln Asn Cys Lys Asn Lys Asn Lys Lys Asn Ile Ala Gln  
1 5 10 15  
Glu Asn Lys Arg Phe Val Glu Ser Thr Lys Lys Met Ala Ser Thr Met  
20 25 30  
Met Thr Thr Leu Pro Gln Phe Asn Gly Leu Arg Ala Thr Lys Ile Ser  
35 40 45  
Ala Ala Pro Val Gln Gly Leu Ala Ser Val Gln Pro Met Arg Arg Lys  
50 55 60  
Gly Asn Gly Ala Leu Gly Ala Lys Cys Asp Phe Ile Gly Ser Ser Thr  
65 70 75 80  
Asn Leu Ile Met Val Thr Ser Thr Thr Leu Met Leu Phe Ala Gly Arg  
85 90 95  
Phe Gly Leu Ala Pro Ser Ala Asn Arg Lys Ala Thr Ala Gly Leu Arg  
100 105 110

Leu Glu Ala Arg Asp Ser Gly Leu Gln Thr Gly Asp Pro Ala Gly Phe  
115 120 125  
Thr Leu Ala Asp Thr Leu Ala Cys Ala Ser Asn Leu Thr Pro Ile Ser  
130 135 140  
Pro Ser Ser Pro Asn Leu Glu Ser Met Lys Thr Ser Pro Met Lys Arg  
145 150 155 160  
Leu Ser Leu Arg Asn Arg Gln Trp Leu Leu Ser Gln Leu Glu Ile Ser  
165 170 175  
Arg Trp Leu Ile  
180

<210> 180  
<211> 153  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..153:Ceres Seq. ID 1386217

<400>180  
Met Ala Ser Thr Met Met Thr Thr Leu Pro Gln Phe Asn Gly Leu Arg  
1 5 10 15  
Ala Thr Lys Ile Ser Ala Ala Pro Val Gln Gly Leu Ala Ser Val Gln  
20 25 30  
Pro Met Arg Arg Lys Gly Asn Gly Ala Leu Gly Ala Lys Cys Asp Phe  
35 40 45  
Ile Gly Ser Ser Thr Asn Leu Ile Met Val Thr Ser Thr Thr Leu Met  
50 55 60  
Leu Phe Ala Gly Arg Phe Gly Leu Ala Pro Ser Ala Asn Arg Lys Ala  
65 70 75 80  
Thr Ala Gly Leu Arg Leu Glu Ala Arg Asp Ser Gly Leu Gln Thr Gly  
85 90 95  
Asp Pro Ala Gly Phe Thr Leu Ala Asp Thr Leu Ala Cys Ala Ser Asn  
100 105 110  
Leu Thr Pro Ile Ser Pro Ser Ser Pro Asn Leu Glu Ser Met Lys Thr  
115 120 125  
Ser Pro Met Lys Arg Leu Ser Leu Arg Asn Arg Gln Trp Leu Leu Ser  
130 135 140  
Gln Leu Glu Ile Ser Arg Trp Leu Ile  
145 150

<210> 181  
<211> 149  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..149:Ceres Seq. ID 1386218

<400>181  
Met Met Thr Thr Leu Pro Gln Phe Asn Gly Leu Arg Ala Thr Lys Ile  
1 5 10 15  
Ser Ala Ala Pro Val Gln Gly Leu Ala Ser Val Gln Pro Met Arg Arg  
20 25 30  
Lys Gly Asn Gly Ala Leu Gly Ala Lys Cys Asp Phe Ile Gly Ser Ser  
35 40 45  
Thr Asn Leu Ile Met Val Thr Ser Thr Thr Leu Met Leu Phe Ala Gly

50 55 60  
Arg Phe Gly Leu Ala Pro Ser Ala Asn Arg Lys Ala Thr Ala Gly Leu  
65 70 75 80  
Arg Leu Glu Ala Arg Asp Ser Gly Leu Gln Thr Gly Asp Pro Ala Gly  
85 90 95  
Phe Thr Leu Ala Asp Thr Leu Ala Cys Ala Ser Asn Leu Thr Pro Ile  
100 105 110  
Ser Pro Ser Ser Pro Asn Leu Glu Ser Met Lys Thr Ser Pro Met Lys  
115 120 125  
Arg Leu Ser Leu Arg Asn Arg Gln Trp Leu Leu Ser Gln Leu Glu Ile  
130 135 140  
Ser Arg Trp Leu Ile  
145

<210> 182  
<211> 65  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..65:Ceres Seq. ID 2025156

<400>182  
Met Ala Leu Val Ala Ala Gly Asn Ile Ser Val Ala Asn Met Asn Lys  
1 5 10 15  
Ala Val Ala Leu Leu Lys Glu Ala Leu Lys Lys Ile Ile Ala Ala  
20 25 30  
Lys Thr Ala Lys Leu Arg Ala Lys Gly Ala Leu Thr Asn Leu Gln Asp  
35 40 45  
Ile Val Ile Arg Gln Ser Lys Val Thr Gly Lys Arg Lys Glu Asp Glu  
50 55 60  
Arg  
65

<210> 183  
<211> 770  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..770:Ceres Seq. ID 1388499

<400>183  
aaaaactca ttctcataac aattccctta actctatctc taggtgaatt aagagtaana  
aattgtagg gcagaggttg aagtagatgt ggaatcaaa tctaoggtgt ataagtttt  
catgtctct agaagatcac aacatgcttc caagccact cgtatgtac aggatgtga  
tctgtcgaa gcgagtgagg gcgaattgg aagcacttc ttgtgaaat taactgtga  
tggagacca aaagtcca aggtatgat tgaaggcatt gatatgaaga tgaatatgat  
ccadggagg gtgttagagg gacctctgaa ggaagagtac aatatctt cgaaacgat  
gaaagtaagc cctaagcaag gaggtctgg aagtgtgtg aagtgaact tgaagtga  
gagaattgat gaaaagtgg ctacactaga gagctcctt cagttctcg tgaattgt  
taagagatt gaccaatcc catgtctga gggtagggg actttctcc ctggttatg  
tgtggagag ttattgatg cgtacgtga aggactctc ttgtgttta tatgtcttc  
tcaataatgt tatctctga gattagaag agaggtacc aagtgctga taccatgt  
ctcaaatgg gaagttgat aacatttaa gtggttatg agttctcat atatagtg  
ttgtgttgg ttgaaatc atattaat gcaacgatac catgttttt

<210> 184  
<211> 151  
<212> PRT  
<213> Arabidopsis thaliana  
<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..151:Ceres Seq. ID 1388500

<400>184  
Met Val Glu Ala Glu Val Glu Val Asp Val Glu Ile Lys Ser Thr Ala  
1 5 10 15  
Asp Lys Phe Met Phe Ser Arg Arg Ser Gln His Ala Ser Lys Ala  
20 25 30  
Thr Arg Tyr Val Gln Gly Cys Asp Leu Leu Glu Gly Thr Gly Glu  
35 40 45  
Val Gly Ser Ile Leu Leu Trp Lys Leu Thr Val Asp Gly Glu Pro Lys  
50 55 60  
Val Ser Lys Asp Met Ile Glu Ala Ile Asp Met Lys Met Asn Met Ile  
65 70 75 80  
Gln Trp Arg Val Leu Glu Gly Pro Leu Lys Glu Tyr Asn Ile Phe  
85 90 95  
Ser Lys Thr Met Lys Val Ser Pro Lys Gln Gly Ser Gly Ser Val  
100 105 110  
Val Lys Trp Asn Leu Lys Tyr Glu Arg Ile Asp Glu Lys Val Ala His  
115 120 125  
Leu Glu Arg Leu Leu Gln Phe Val Glu Cys Val Asn Glu Ile Asp  
130 135 140  
Gln Tyr Leu Leu Ser Glu Gly  
145 150

<210> 185  
<211> 131  
<212> PRT  
<213> Arabidopsis thaliana  
<220>  
<223> any n or Xaa = unknown  
<223> LOCATION: 1..131:Ceres Seq. ID 1388501

<400>185  
Met Phe Ser Arg Arg Ser Gln His Ala Ser Lys Ala Thr Arg Tyr Val  
1 5 10 15  
Gln Gly Cys Asp Leu Leu Glu Gly Glu Trp Gly Glu Val Gly Ser Ile  
20 25 30  
Leu Leu Trp Lys Leu Thr Val Asp Gly Glu Pro Lys Val Ser Lys Asp  
35 40 45  
Met Ile Glu Ala Ile Asp Met Lys Met Asn Met Ile Gln Trp Arg Val  
50 55 60  
Leu Glu Gly Pro Leu Lys Glu Glu Tyr Asn Ile Phe Ser Lys Thr Met  
65 70 75 80  
Lys Val Ser Pro Lys Gln Gly Gly Ser Gly Ser Val Val Lys Trp Asn  
85 90 95  
Leu Lys Tyr Glu Arg Ile Asp Glu Lys Val Ala His Leu Glu Arg Leu  
100 105 110  
Leu Gln Phe Phe Val Glu Cys Val Asn Glu Ile Asp Gln Tyr Leu Leu  
115 120 125  
Ser Glu Gly  
130

<210> 186  
<211> 83  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..83:Ceres Seq. ID 1388502

<400>186  
Met Ile Glu Ala Ile Asp Met Lys Met Asn Met Ile Gln Trp Arg Val  
1 5 10 15  
Leu Glu Gly Pro Leu Lys Glu Tyr Asn Ile Phe Ser Lys Thr Met  
20 25 30  
Lys Val Ser Pro Lys Gln Gly Gly Ser Gly Ser Val Val Lys Trp Asn  
35 40 45  
Leu Lys Tyr Glu Arg Ile Asp Glu Lys Val Ala His Leu Glu Arg Leu  
50 55 60  
Leu Gln Phe Phe Val Glu Cys Val Asn Glu Ile Asp Gln Tyr Leu Leu  
65 70 75 80  
Ser Glu Gly

<210> 187  
<211> 752  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..752:Ceres Seq. ID 1388519

<400>187  
aatcacgc gcttcatca tcttcttctt agatcgctc ctccactatt cgaaaaatg  
60  
gcgaactgg tgatgtgtt ggttcttgt atcttacgg cgatagccat ggaggaaagg  
120  
agggaataa ttggaagaa tacaatggtg gtccaagta gcacctactg tgacattgc  
180  
aaattcgct tcgagatcc tgaattctcc tacttcatcc ccggtgcaac ggtgaagta  
240  
tcagtcaag acaggaagac aatggaagag gtttacacag acaaatctot atcagacaa  
300  
gaaggaaagt ataatgtcat tgttcacgac gatcacagag accagatgtg cgatgtttg  
360  
cttggaaaa gtcggataa aaactgtctt aaatctccg ttggagctga gaagtctgt  
420  
gtgatctga accattacag tggcatggc tcgcagatca gacatgtaa caacatgga  
480  
ttggaagaag agtgagtga tgtgtctgc tctgctttgt ttccagaagta tatggtgat  
540  
gaagatgagg atgataataa aaacctctc taatctctct gtttaactct atgactgtt  
600  
gtttcttca ttaagtgtt tccagttatg gaagagatat atttgtattt gtttgattac  
660  
ttattgttg tttttagaga ttttgactct ggtgatcgga taactatctg tttgtgtaag  
720  
cttcattat attgatgtgt catttcccttg cc  
752

<210> 188  
<211> 190  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..190:Ceres Seq. ID 1388520

<400>188

Asn Leu Thr Arg Phe His Leu Leu Pro Arg Phe Ala Pro Pro Leu  
 1 5 10 15  
 Phe Glu Lys Met Ala Lys Leu Val Met Leu Val Leu Cys Ile Leu  
 20 25 30  
 Pro Ala Ile Ala Met Ala Ala Arg Arg Gly Asn Ile Gly Lys Asn Thr  
 35 40 45  
 Met Val Val Gln Gly Ser Thr Tyr Cys Asp Ile Cys Lys Phe Gly Phe  
 50 55 60  
 Glu Thr Pro Glu Ser Ser Tyr Phe Ile Pro Gly Ala Thr Val Lys Leu  
 65 70 75 80  
 Ser Cys Lys Asp Arg Lys Thr Met Glu Glu Val Tyr Thr Asp Lys Ala  
 85 90 95  
 Val Ser Asp Lys Glu Gly Lys Tyr Lys Phe Ile Val His Asp Asp His  
 100 105 110  
 Arg Asp Gln Met Cys Asp Val Leu Leu Val Lys Ser Ser Asp Lys Thr  
 115 120 125  
 Cys Ser Lys Ile Ser Val Gly Arg Glu Lys Ser Arg Val Ile Leu Asn  
 130 135 140  
 His Tyr Ser Gly Ile Ala Ser Gln Ile Arg His Ala Asn Asn Met Gly  
 145 150 155  
 Phe Glu Lys Glu Val Ser Asp Val Phe Cys Ser Ala Leu Phe Gln Lys  
 165 170 175  
 Tyr Met Val Asp Glu Asp Glu Asp Ile Lys Asn His Leu  
 180 185 190

&lt;210&gt; 189

&lt;211&gt; 171

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..171:Ceres Seq. ID 1388521

<400>189  
 Met Ala Lys Leu Val Met Leu Leu Val Leu Cys Ile Leu Pro Ala Ile  
 1 5 10 15  
 Ala Met Ala Ala Arg Arg Gly Asn Ile Gly Lys Asn Thr Met Val Val  
 20 25 30  
 Gln Gly Ser Thr Tyr Cys Asp Ile Cys Lys Phe Gly Phe Glu Thr Pro  
 35 40 45  
 Glu Ser Ser Tyr Phe Ile Pro Gly Ala Thr Val Lys Leu Ser Cys Lys  
 50 55 60  
 Asp Arg Lys Thr Met Glu Glu Val Tyr Thr Asp Lys Ala Val Ser Asp  
 65 70 75 80  
 Lys Glu Gly Lys Tyr Lys Phe Ile Val His Asp Asp His Arg Asp Gln  
 85 90 95  
 Met Cys Asp Val Leu Leu Val Lys Ser Ser Asp Lys Thr Cys Ser Lys  
 100 105 110  
 Ile Ser Val Gly Arg Glu Lys Ser Arg Val Ile Leu Asn His Tyr Ser  
 115 120 125  
 Gly Ile Ala Ser Gln Ile Arg His Ala Asn Asn Met Gly Phe Glu Lys  
 130 135 140  
 Glu Val Ser Asp Val Phe Cys Ser Ala Leu Phe Gln Lys Tyr Met Val  
 145 150 155  
 Asp Glu Asp Glu Asp Asp Ile Lys Asn His Leu  
 165 170 175

&lt;210&gt; 190

&lt;211&gt; 166

<212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..166:Ceres Seq. ID 1388522  
 <400>190  
 Met Leu Leu Val Leu Cys Ile Leu Pro Ala Ile Ala Met Ala Ala Arg  
 1 5 10 15  
 Arg Gly Asn Ile Gly Lys Asn Thr Met Val Val Gln Gly Ser Thr Tyr  
 20 25 30  
 Cys Asp Ile Cys Lys Phe Gly Phe Glu Thr Pro Glu Ser Ser Tyr Phe  
 35 40 45  
 Ile Pro Gly Ala Thr Val Lys Leu Ser Cys Lys Asp Arg Lys Thr Met  
 50 55 60  
 Glu Glu Val Tyr Thr Asp Lys Ala Val Ser Asp Lys Glu Gly Lys Tyr  
 65 70 75 80  
 Lys Phe Ile Val His Asp Asp His Arg Asp Gln Met Cys Asp Val Leu  
 85 90 95  
 Leu Val Lys Ser Ser Asp Lys Thr Cys Ser Lys Ile Ser Val Gly Arg  
 100 105 110  
 Glu Lys Ser Arg Val Ile Leu Asn His Tyr Ser Gly Ile Ala Ser Gln  
 115 120 125  
 Ile Arg His Ala Asn Asn Met Gly Phe Glu Lys Glu Val Ser Asp Val  
 130 135 140  
 Phe Cys Ser Ala Leu Phe Gln Lys Tyr Met Val Asp Glu Asp Glu Lys  
 145 150 155  
 Asp Ile Lys Asn His Leu  
 165 166

&lt;210&gt; 191

&lt;211&gt; 658

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..658:Ceres Seq. ID 1388563

<400>191  
 aagaagttt cttacatcc acaaaagaaa aaagtttcc caaaaaacaa aaacttacac  
 agtttagat atggatcttg attactaca agattgttc aaattcaatt tcccaacac  
 catcaagatc cgtaccacaaa cttcaaaaga caaagaaga cgaagtgat gagcagaag  
 atgacctcag ctgcagcaca cccacatccc aagaacacaa gattccgcc gtacagact  
 ctccacctcg tccgcgaga aaaccgcgc caccacgtc agcacgtcg gctacgcggg  
 cctcgatgat cagatctgc aaggaagc ttttagtgc gactgtgag ataactaga  
 atcgggaaga gattgacct tcttctct cgtctaca tggacgtcg actacggcta  
 aacgcggag aggttacct tatgtctc gaagatgagg cttaactcaa tattacatt  
 tttttacgt ttactcgaa atattgaa ataatattc tgttggttt cgggtttaa  
 tattttaat ttatgatg atattgtg taattttct gcaaccgc atattaatt  
 cgcatggagg ggtcgatgtt gtaaatgag taataatga aatttaattc caatgccc  
 60 120 180 240 300 360 420 480 540 600 658  
 <210> 192  
 <211> 55  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..55:Ceres Seq. ID 1388564

<400>192

Arg Ser Phe Leu Tyr Ile His Lys Lys Arg Lys Lys Val Ser Gln Lys Thr  
1 5 10 15  
Lys Thr Tyr Thr Ser Leu Asp Met Asp Leu Glu Leu Leu Gln Asp Leu  
20 25 30  
Ser Lys Phe Asn Phe Pro Thr Pro Ile Lys Ile Arg Ser Lys Thr Ser  
35 40 45  
Lys Asp Lys Glu Gly Arg Arg  
50 55

<210> 193

<211> 50

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..50:Ceres Seq. ID 1388565

<400>193

Met Ile Arg Ser Cys Lys Arg Lys Leu Val Ser Thr Cys Glu Ile  
1 5 10 15  
Ile Met Asn Arg Glu Glu Ile Asp Arg Phe Phe Ser Ser Val Tyr Asn  
20 25 30  
Glu Thr Ser Thr Thr Ala Lys Arg Arg Arg Ser Tyr Pro Tyr Cys Ser  
35 40 45  
Arg Arg  
50

<210> 194

<211> 39

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..39:Ceres Seq. ID 1388566

<400>194

Met Arg Leu Asn Ser Ile Phe Thr Phe Phe Tyr Ser Phe Thr Gly Thr  
1 5 10 15  
Ile Val Lys Leu Ile Ile Cys Trp Cys Ser Val Leu Asn Ile Phe Asn  
20 25 30  
Leu Val Met Asn Met Asp Gly  
35

<210> 195

<211> 866

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..866:Ceres Seq. ID 1388793

<400>195

aacgggtgg gtaataetca cgcgctttca tcatctctt cctagattcg ctcctccact  
gtcttaatt tcatatctt ttcttcatt ttcttcctt ctcactcaa atccataatc  
gtctactc tctctctc tagattcgaa aaatggcg aactggtgat gtgttggtt  
ctcttatct taccggcat accatggcg gaaaggagg gaaatattgg aagaataca  
atgggtgtc aagtagcac ctactgac attgcaat tggcttoga gactctgaa  
tctctact tctcccggtg tgaacgggtg aagctatcat gcaagacag gaagacaatg  
gaagaggtt acacagaca accgtgtatc gacaagaag gaaagtataa gtccattgc  
tcacagcat acagagaca gatgtgat gtttgttg tgaagctc ggataaacc  
tgcttaaaa tctcgttg acgtgaga tctgtgtg tcttgacca ttcagtggc  
atgctcgc acatcagaca tctcaaca atggattg agaagaagt gactgatg  
ttctgtctg ctgtgttca gaagtatat gtgatgaag atgagatga tatteaaac  
catctaat ctctctgtt aatctatga tctgtgtt ttctcataa tgagtttga  
gttatggaag agatatatt gtattgtt gattactat ttgttctt tagagattg  
gactctgtg atcgataac tatctgtt tgaagctc ttatatattg atatgtcatt  
tcttgctta ggaagtgtc ttatt

<210> 196

<211> 171

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..171:Ceres Seq. ID 1388794

<400>196

Met Ala Lys Leu Val Met Leu Leu Val Leu Cys Ile Leu Pro Ala Ile  
1 5 10 15  
Ala Met Ala Ala Arg Arg Gly Asn Ile Gly Lys Asn Thr Met Val Val  
20 25 30  
Gln Gly Ser Thr Tyr Cys Asp Ile Cys Lys Phe Gly Phe Glu Thr Pro  
35 40 45  
Glu Ser Ser Tyr Phe Ile Pro Gly Ala Thr Val Lys Leu Ser Cys Lys  
50 55 60  
Asp Arg Lys Thr Met Glu Glu Val Tyr Thr Asp Lys Ala Val Ser Asp  
65 70 75 80  
Lys Glu Gly Lys Tyr Lys Phe Ile Val His Asp Asp His Arg Asp Gln  
85 90 95  
Met Cys Asp Val Leu Leu Val Lys Ser Ser Asp Lys Thr Cys Ser Lys  
100 105 110  
Ile Ser Val Gly Arg Glu Lys Ser Arg Val Ile Leu Asn His Tyr Ser  
115 120 125  
Gly Ile Ala Ser Gln Ile Arg His Ala Asn Asn Met Gly Phe Glu Lys  
130 135 140  
Glu Val Ser Asp Val Phe Cys Ser Ala Leu Phe Gln Lys Tyr Met Val  
145 150 155 160  
Asp Glu Asp Glu Asp Asp Ile Lys Asn His Leu  
165 170

<210> 197

<211> 166

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..166:Ceres Seq. ID 1388795



<400>197  
Met Leu Leu Val Leu Cys Ile Leu Pro Ala Ile Ala Met Ala Ala Arg  
1 5 10 15  
Arg Gly Asn Ile Gly Lys Asn Thr Met Val Val Gln Gly Ser Thr Tyr  
20 25 30  
Cys Asp Ile Cys Lys Phe Gly Phe Glu Thr Pro Glu Ser Ser Tyr Phe  
35 40 45  
Ile Pro Gly Ala Thr Val Lys Leu Ser Cys Lys Asp Arg Lys Thr Met  
50 55 60  
Glu Glu Val Tyr Thr Asp Lys Ala Val Ser Asp Lys Glu Gly Lys Tyr  
65 70 75 80  
Lys Phe Ile Val His Asp Asp His Arg Asp Gln Met Cys Asp Val Leu  
85 90 95  
Leu Val Lys Ser Ser Asp Lys Thr Cys Ser Lys Ile Ser Val Gly Arg  
100 105 110  
Glu Lys Ser Arg Val Ile Leu Asn His Tyr Ser Gly Ile Ala Ser Gln  
115 120 125  
Ile Arg His Ala Asn Asn Met Gly Phe Glu Lys Glu Val Ser Asp Val  
130 135 140  
Phe Cys Ser Ala Leu Phe Gln Lys Tyr Met Val Asp Glu Asp Glu Asp  
145 150 155 160  
Asp Ile Lys Asn His Leu  
165

<210> 198  
<211> 154  
<212> PRT  
<213> Arabidopsis thaliana  
<220>  
<223> any n or Xaa = unknown  
<223> LOCATION: 1..154:Ceres Seq. ID 1388796

<400>198  
Met Ala Ala Arg Arg Gly Asn Ile Gly Lys Asn Thr Met Val Val Gln  
1 5 10 15  
Gly Ser Thr Tyr Cys Asp Ile Cys Lys Phe Gly Phe Glu Thr Pro Glu  
20 25 30  
Ser Ser Tyr Phe Ile Pro Gly Ala Thr Val Lys Leu Ser Cys Lys Asp  
35 40 45  
Arg Lys Thr Met Glu Glu Val Tyr Thr Asp Lys Ala Val Ser Asp Lys  
50 55 60  
Glu Gly Lys Tyr Lys Phe Ile Val His Asp Asp His Arg Asp Gln Met  
65 70 75 80  
Cys Asp Val Leu Leu Val Lys Ser Ser Asp Lys Thr Cys Ser Lys Ile  
85 90 95  
Ser Val Gly Arg Glu Lys Ser Arg Val Ile Leu Asn His Tyr Ser Gly  
100 105 110  
Ile Ala Ser Gln Ile Arg His Ala Asn Asn Met Gly Phe Glu Lys Glu  
115 120 125  
Val Ser Asp Val Phe Cys Ser Ala Leu Phe Gln Lys Tyr Met Val Asp  
130 135 140  
Glu Asp Glu Asp Asp Ile Lys Asn His Leu  
145 150

<210> 199  
<211> 533  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown  
<223> LOCATION: 1..533:Ceres Seq. ID 1392041  
<400>199  
aacataaca caatacaacc aaaaaaaaaa aaacagaaaag taacatgaga tctctcttat  
tagcgtgty cctgtttctt gctttacct gcgtgaagc agcgtgtct tgaacaacgg  
tgattggga tctttacct tgtttacct acgtgactca ggcgcgaacg gtccaaccc  
tctgtcgaa cgtgtcaca acactcaaga gtcaggctca aacttctgt gacgtcagg  
300 340 380 420  
gggtctcgt ttgcatacna tctgtctatg gaggaactac tctctctctt agaacctcc  
aaatgtttt ggaattgctt tctaaatgtg gtgtgatct cctttacaag ttacgcctt  
cactgactg agcagatcag cagtgagaca agcagaaaat cttaaaggaa gctactcaa  
420 460 500 540  
gaataaat acactaatna ttaataatg agggcatgg ttgttagtt gtaattgat  
cagtgatga ttgtctattt gaattgtcta atatcagcag gcaattatct ctg

<210> 200  
<211> 113  
<212> PRT  
<213> Arabidopsis thaliana  
<220>  
<223> any n or Xaa = unknown  
<223> LOCATION: 1..113:Ceres Seq. ID 1392042

<400>200  
Met Arg Ser Leu Leu Leu Ala Val Cys Leu Val Leu Ala Leu His Cys  
1 5 10 15  
Gly Glu Ala Ala Val Ser Cys Asn Thr Val Ile Ala Asp Leu Tyr Pro  
20 25 30  
Cys Leu Ser Tyr Val Thr Gln Gly Gly Pro Val Pro Thr Leu Cys Cys  
35 40 45  
Asn Gly Leu Thr Thr Leu Lys Ser Gln Ala Gln Thr Ser Val Asp Arg  
50 55 60  
Gln Gly Val Cys Arg Cys Ile Lys Ser Ala Ile Gly Gly Leu Thr Leu  
65 70 75 80  
Ser Pro Arg Thr Ile Gln Asn Ala Leu Glu Leu Pro Ser Lys Cys Gly  
85 90 95  
Val Asp Leu Pro Tyr Lys Phe Ser Pro Ser Thr Asp Cys Asp Ser Ile  
100 105 110  
Gln

<210> 201  
<211> 1028  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown  
<223> LOCATION: 1..1028:Ceres Seq. ID 1393342

<400>201  
aatatgtctt taattgtctt tatcgagaa ggcttactct gttttcagg cgaactctgt  
ttattctctc cgagctcgta cgagaatatt cgtaaaggtt tcttgagc ttaacagtg  
agtaccttg agaccacaag agctgagctt ggctgtgtag ttgttattt gaataagcg  
300 340 380 420  
gggcgagag acaagatttg tcgagattt caatgtggtt ccaagtctt gactgtggc  
caactagca ctgctcaaaa tgtcgacaa aacaccagct tggctaggaa agtttccgt

ctcttaagt ttgtaaatga tcttcatgt ctcattagcc ctgttcccaa agggactcca 360  
 ctcgcccttg tctgtccgg aagctctaaa aacgcgttgc tctcaagtt ctgttctcta 420  
 gatacaatt tctgcttgg cagactggg atttacaagg acaagaacg tctgtagatt 480  
 ctgtgaaga tatacattt ctgttggatg gtttcttgg ttgtcacatc ctgttgtag 540  
 gtggggagc ttgttagct gtcagatca attaagaagt tagaanaaga gatagggaac 600  
 agggataac acagaatgt gcaatacgt gcaaaagtat agaatacaa cgaagagtea 660  
 ttggctctga tcaagacg gatgatgta gttgtgctt tggattgct tcaattgset 720  
 caaagaag taatcccg agtcacagt gctttggat tgcctcatc acttatctc 780  
 ttgtatcagc tatcgctc acatccaag tccaagatg tctgataaa gggaaacaag 840  
 gaaggcga ctagttcat tgtccacag tccatatat agtttgtc tgttttacta 900  
 ttaactgtg ttaattgt ctgtctgac tacttaatg ttgtgttag attcaggt 960  
 atataagta atgataaaca ttgcagcttc ttagaataa taagaacsga taagcattac 1020  
 gagggtgc 1028

<210> 202

<211> 235

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..235:Ceres Seq. ID 1393343

<400>202

Met Ser Thr Leu Glu Thr Thr Arg Ala Glu Leu Gly Leu Val Val Val 15  
 1 5 10  
 Tyr Leu Asn Lys Ala Glu Ala Arg Asp Lys Ile Cys Arg Ala Ile Gln 30  
 20 25  
 Tyr Gly Ser Lys Phe Leu Ser Asp Gly Gln Pro Gly Thr Ala Gln Asn 45  
 35 40  
 Val Asp Lys Asn Thr Ser Leu Ala Arg Lys Val Phe Arg Leu Phe Lys 60  
 50 55  
 Phe Val Asn Asp Leu His Ala Leu Ile Ser Pro Val Pro Lys Gly Thr 80  
 65 70 75  
 Pro Leu Pro Leu Val Leu Leu Gly Lys Ser Lys Asn Ala Leu Leu Ser 95  
 85 90  
 Thr Phe Leu Phe Leu Asp Gln Ile Val Trp Leu Gly Arg Thr Gly Ile 110  
 100 105  
 Tyr Lys Asp Lys Glu Arg Ala Glu Ile Leu Gly Arg Ile Ser Leu Phe 125  
 115 120  
 Cys Trp Met Gly Ser Ser Val Cys Thr Ser Leu Val Glu Val Gly Glu 140  
 130 135  
 Leu Gly Arg Leu Ser Ala Ser Ile Lys Lys Leu Glu Lys Glu Ile Gly 160  
 145 150 155  
 Asn Lys Asp Lys His Gln Asn Glu Gln Tyr Arg Ala Lys Val Glu Lys 175  
 165 170  
 Ser Asn Glu Arg Ser Leu Ala Leu Ile Lys Ala Gly Met Asp Val Val 190  
 180 185  
 Val Ala Phe Gly Leu Leu Gln Leu Ala Pro Lys Lys Val Thr Pro Arg 205  
 195 200  
 Val Thr Gly Ala Phe Gly Phe Ala Ser Ser Leu Ile Ser Cys Tyr Gln 220  
 210 215  
 Leu Leu Pro Ser His Pro Lys Ser Lys Met Val 235  
 225 230

<210> 203

<211> 653

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..146:Ceres Seq. ID 1393558

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..653:Ceres Seq. ID 1393556

<400>203

tcttccaat ccaacagct tcatgtttt atcttccaat caaagcagct caaatcaagc 60  
 tgcgaagat gatgacatg accttactt ggggaatcaa agccaatct ctcttgatt 120  
 tctgaaaac tgactcatg cttagtata tctcaattt aatcacttt aatcgttgc ttgtttct 180  
 ccgtttcta tcaatactc gagaatgcc gaatccaatt caaatccctt tcttctccc 240  
 cgtgtctcc tcaacgcct cgtcttctt ccgggtctc cggcctctt atccataat 300  
 gcgtaccag atccgcgct aagctgctt cggttctct ttctggctc aacgcaga 360  
 tctgtactt gctgatgctt gcagctatg ctttcaacg aggtgttct atcgcattg 420  
 tctgcgatt aacgcgga tacygtgtt tttagatga tgaaggcgtt gctgatacgg 480  
 ccacgatga tccatgcca tgtgttgat aatgatgat aataaatca aatcggaaa 540  
 aatcctaaa aaggtgtt ttataatga tttgttatt gttgttgtt tctgtatg 600  
 tagttgctt gtaacaaa aatccaatct aatgtgttt cttgagtata tgc 653

<210> 204

<211> 168

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..168:Ceres Seq. ID 1393557

<400>204

Phe Pro Ile Pro Thr Ala Ser Ser Phe Tyr Leu Pro Ile Lys Ala Ala 15  
 1 5 10  
 Gln Ile Lys Leu Ala Lys Met Met His Met Thr Phe Tyr Trp Gly Ile 30  
 20 25  
 Lys Ala Thr Ile Leu Phe Asp Phe Trp Lys Thr Asp Ser Trp Leu Ser 45  
 35 40  
 Tyr Ile Leu Thr Leu Ile Ala Cys Phe Val Phe Ser Ala Phe Tyr Gln 60  
 50 55  
 Tyr Leu Glu Asn Arg Arg Ile Gln Phe Lys Ser Leu Ser Ser Ser Arg 80  
 65 70 75  
 Arg Ala Pro Pro Pro Arg Ser Ser Ser Gly Val Ser Ala Pro Leu 95  
 85 90  
 Ile Pro Lys Ser Gly Thr Arg Ser Ala Ala Lys Ala Ala Ser Val Leu 110  
 100 105  
 Leu Phe Gly Val Asn Ala Ala Ile Gly Tyr Leu Leu Met Leu Ala Ala 125  
 115 120  
 Met Ser Phe Asn Gly Gly Val Phe Ile Ala Ile Val Val Gly Leu Thr 140  
 130 135  
 Ala Gly Tyr Ala Val Phe Arg Ser Asp Asp Gly Ala Asp Thr Ala 160  
 145 150 155  
 Thr Asp Asp Pro Cys Pro Cys Ala 165

<210> 205

<211> 146

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..146:Ceres Seq. ID 1393558

<400>205  
Met Met His Met Thr Phe Tyr Trp Gly Ile Lys Ala Thr Ile Leu Phe  
1 5 10 15  
Asp Phe Trp Lys Thr Asp Ser Trp Leu Ser Tyr Ile Leu Thr Leu Ile  
20 25 30  
Ala Cys Phe Val Phe Ser Ala Phe Tyr Gln Tyr Leu Glu Asn Arg Arg  
35 40 45  
Ile Gln Phe Lys Ser Leu Ser Ser Arg Arg Ala Pro Pro Pro  
50 55 60  
Arg Ser Ser Ser Gly Val Ser Ala Pro Leu Ile Pro Lys Ser Gly Thr  
65 70 75 80  
Arg Ser Ala Ala Lys Ala Ala Ser Val Leu Leu Phe Gly Val Asn Ala  
85 90 95  
Ala Ile Gly Tyr Leu Leu Met Leu Ala Ala Met Ser Phe Asn Gly Gly  
100 105 110  
Val Phe Ile Ala Ile Val Gly Leu Thr Ala Gly Tyr Ala Val Phe  
115 120 125  
Arg Ser Asp Asp Gly Gly Ala Asp Thr Ala Thr Asp Asp Pro Cys Pro  
130 135 140  
Cys Ala  
145

<210> 206  
<211> 145  
<212> PRT  
<213> Arabidopsis thaliana  
  
<220>  
<223> any n or Xaa = unknown  
  
<223> LOCATION: 1..145:Ceres Seq. ID 1393559

<400>206  
Met His Met Thr Phe Tyr Trp Gly Ile Lys Ala Thr Ile Leu Phe Asp  
1 5 10 15  
Phe Trp Lys Thr Asp Ser Trp Leu Ser Tyr Ile Leu Thr Leu Ile Ala  
20 25 30  
Cys Phe Val Phe Ser Ala Phe Tyr Gln Tyr Leu Glu Asn Arg Ile  
35 40 45  
Gln Phe Lys Ser Leu Ser Ser Arg Arg Ala Pro Pro Pro Pro Arg  
50 55 60  
Ser Ser Ser Gly Val Ser Ala Pro Leu Ile Pro Lys Ser Gly Thr Arg  
65 70 75 80  
Ser Ala Ala Lys Ala Ala Ser Val Leu Leu Phe Gly Val Asn Ala Ala  
85 90 95  
Ile Gly Tyr Leu Leu Met Leu Ala Ala Met Ser Phe Asn Gly Gly Val  
100 105 110  
Phe Ile Ala Ile Val Val Gly Leu Thr Ala Gly Tyr Ala Val Phe Arg  
115 120 125  
Ser Asp Asp Gly Gly Ala Asp Thr Ala Thr Asp Asp Pro Cys  
130 135 140  
Ala  
145

<210> 207  
<211> 525  
<212> DNA  
<213> Arabidopsis thaliana  
  
<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..525:Ceres Seq. ID 1396782

<400>207  
atgatttta cagacgagag gagcggcga caatgcgaa ttgatgatg agattacca  
120  
ttagcttgag aagcttctct gttcagctt ctatcccaa cggctgcgcg aggtgatcg  
180  
gaggtatcag cggcgggtga ggaacgatga ttggaatt accgttgag agatacga  
240  
gacgttgat ggaacgaatg tccaacgatc agaacaagt gaaagatg attactatg  
300  
tcgtcaaat cggcttctca gttcggatg atgcatga agttgatga acttactatg  
360  
ggttcggg atgtcacaga taccagcgc atcagaagt aggtttcca acaatcgtt  
420  
gcaaatcgc taaggaaac aaggaaatc taaggatca tcttcgctga agaggtttt  
480  
actttatgt atattgttg tatgataga actcgatat gtgatgaa gtaagaata  
525  
agaactttgt acaactttgt gatagattta gtacgggat tctgc

<210> 208  
<211> 135  
<212> PRT  
<213> Arabidopsis thaliana  
  
<220>  
<223> any n or Xaa = unknown  
  
<223> LOCATION: 1..135:Ceres Seq. ID 1396783

<400>208  
Asp Phe Tyr Arg Arg Glu Glu Thr Ala Thr Met Ala Asn Leu Met Met  
1 5 10 15  
Arg Leu Pro Ile Ser Leu Arg Ser Phe Ser Val Ser Ala Ser Ser Ser  
20 25 30  
Asn Gly Ser Pro Pro Val Ile Gly Gly Ser Ser Gly Val Gly Pro  
35 40 45  
Met Ile Val Glu Leu Pro Leu Glu Lys Ile Arg Arg Pro Leu Met Arg  
50 55 60  
Thr Arg Ser Asn Asp Gln Asn Lys Val Lys Glu Leu Met Asp Ser Ile  
65 70 75 80  
Arg Gln Ile Gly Leu Gln Val Pro Ile Asp Val Ile Glu Val Asp Gly  
85 90 95  
Thr Tyr Tyr Gly Phe Ser Gly Cys His Arg Tyr Glu Ala His Gln Lys  
100 105 110  
Leu Gly Leu Pro Thr Ile Arg Cys Lys Ile Arg Lys Gly Thr Lys Glu  
115 120 125  
Thr Leu Arg His His Leu Arg  
130 135

<210> 209  
<211> 125  
<212> PRT  
<213> Arabidopsis thaliana  
  
<220>  
<223> any n or Xaa = unknown  
  
<223> LOCATION: 1..125:Ceres Seq. ID 1396784

<400>209  
Met Ala Asn Leu Met Met Arg Leu Pro Ile Ser Leu Arg Ser Phe Ser  
1 5 10 15  
Val Ser Ala Ser Ser Ser Asn Gly Ser Pro Pro Val Ile Gly Gly Ser  
20 25 30  
Ser Gly Gly Val Gly Pro Met Ile Val Glu Leu Pro Leu Glu Lys Ile

35  
Arg Arg Pro Leu Met Arg Thr Arg Ser Asn Asp Gln Asn Lys Val Lys 45  
50  
Glu Leu Met Asp Ser Ile Arg Gln Ile Gly Leu Gln Val Pro Ile Asp 60  
65 70 75  
Val Ile Glu Val Asp Gly Thr Tyr Tyr Gly Phe Ser Gly Cys His Arg 80  
85  
Tyr Glu Ala His Gln Lys Leu Gly Leu Pro Thr Ile Arg Cys Lys Ile 95  
100 105  
Arg Lys Gly Thr Lys Glu Thr Leu Arg His His Leu Arg 110  
115 120 125

&lt;210&gt; 210

&lt;211&gt; 121

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..121:Ceres Seq. ID 1396785

<400>210  
Met Met Arg Leu Pro Ile Ser Leu Arg Ser Phe Ser Val Ser Ala Ser  
1 5 10 15  
Ser Ser Asn Gly Ser Pro Val Ile Gly Gly Ser Ser Gly Val  
20 25 30  
Gly Pro Met Ile Val Glu Leu Pro Leu Glu Lys Ile Arg Arg Pro Leu  
35 40 45  
Met Arg Thr Arg Ser Asn Asp Gln Asn Lys Val Lys Glu Leu Met Asp  
50 55 60  
Ser Ile Arg Gln Ile Gly Leu Gln Val Pro Ile Asp Val Ile Glu Val  
65 70 75 80  
Asp Gly Thr Tyr Tyr Gly Phe Ser Gly Cys His Arg Tyr Glu Ala His  
85 90 95  
Gln Lys Leu Gly Leu Pro Thr Ile Arg Cys Lys Ile Arg Lys Gly Thr  
100 105 110  
Lys Glu Thr Leu Arg His His Leu Arg  
115 120

&lt;210&gt; 211

&lt;211&gt; 939

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..939:Ceres Seq. ID 1396802

<400>211  
atacaaac aggtcagg tttatcttc catattctt toccactgat tctacaatc 60  
gatcaagg aatcacgaa ctgttggta attctggty ggaagatct tgatttgc 120  
gttggtyt gaaatagg gagtctgt ttgtgttg gtcaccggt gatgactct 180  
tcgacagtc catgtttatg tctggatgt cacacagcg gtaccatcag gaggaaga 240  
atcttggca aagcaagaa tctggaaata ggaagctott tcactggctc agcatgtt 300  
tttaggcit ctcgaagcg tgcacgaga attggaatc gaagagcaa aagctctt 360  
atttgaat aggatgtgc tggatattac gacatactc tgggtgatg acaaaagcaa 420  
attgtaat atttaccata caaagctgtg aggaagctc ttcacagct gtacgaagt 480  
aatctctc aatcacttg gtctataat catatcataa caaacagacc gacgatggc 540  
aaacgttcc tccgtgcct cggcaaggag agtcaggagc ttgcagaaag agtgatgatc 600

acgctcttc actgtatgg caaatgac aagaatgcg accatgggaa gatatacag 660  
gaatacagg atgaacctt ggcgttgatg cgtgaacgc tgatggagc cgtgatatgg 720  
ctctcagtg acaaaactc aagtaaatg igtgaagaag acagcgaca ccacagaca 780  
acaattgctt tgcacattg cttctcagtt tttaggactt cgagttttt agttgatta 840  
catattacta tagtgaaga gcagtgtga tatgatgct attcttgcca tgcactat 900  
gatattagg gcaaaaaa acaccttct tgtctcct 939

&lt;210&gt; 212

&lt;211&gt; 248

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..248:Ceres Seq. ID 1396803

<400>212  
Ile Thr Asn Gln Gly Gln Gly Phe Ile Phe His Ile Leu Phe Pro Leu  
1 5 10 15  
Ile Leu Thr Ile Asp Leu Arg Glu Ser Thr Asn Cys Leu Val Ile Ser  
20 25 30  
Gly Gly Lys Ser Leu Asp Leu Ala Val Trp Phe Gly Val Met Val Ser  
35 40 45  
Ala Trp Phe Val Val Gly Ser Pro Val Met Asp Ser Ser Ser Pro  
50 55 60  
Cys Leu Cys Leu Asp Ala His Thr Thr Gly Thr Ile Arg Arg Lys  
65 70 75 80  
Ile Leu Gly Lys Ala Arg Asn Leu Glu Leu Gly Ser Ser Phe Thr Gly  
85 90 95  
Ser Arg Ile Val Phe Arg Leu Ser Pro Lys Arg Val Ser Arg Ile Ala  
100 105 110  
Asn Arg Lys Ser Lys Lys Leu Leu Ile Val Asn Glu Asp Val Ala Gly  
115 120  
Asn Tyr Asp Asp Thr Phe Gly Asp Val Gln Lys Gln Ile Val Asn Tyr  
130 135 140  
Phe Thr Tyr Lys Ala Val Arg Thr Val Leu His Gln Leu Tyr Glu Met  
145 150 155 160  
Asn Pro Pro Gln Tyr Thr Trp Phe Tyr Asn His Ile Ile Thr Asn Arg  
165 170 175  
Pro Thr Asp Gly Lys Arg Phe Leu Arg Ala Leu Gly Lys Glu Ser Gln  
180 185 190  
Glu Leu Ala Glu Arg Val Met Ile Thr Arg Leu His Leu Tyr Gly Lys  
195 200 205  
Trp Ile Lys Lys Cys Asp His Gly Lys Ile Tyr Gln Glu Ile Ser Asp  
210 215 220  
Glu Asn Leu Ala Leu Met Arg Glu Arg Leu Met Glu Thr Val Ile Trp  
225 230 235 240  
Pro Ser Asp Asp Thr Asn Ser Arg 245

&lt;210&gt; 213

&lt;211&gt; 203

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..203:Ceres Seq. ID 1396804

<400>213  
Met Val Ser Ala Trp Phe Val Val Gly Ser Pro Val Met Asp Ser Ser  
1 5 10 15  
Ser Ser Pro Cys Leu Cys Leu Asp Ala His Thr Thr Gly Thr Ile Arg  
20 25 30  
Arg Lys Lys Ile Leu Gly Lys Ala Arg Asn Leu Glu Leu Gly Ser Ser  
35 40 45  
Phe Thr Gly Ser Arg Ile Val Phe Arg Leu Ser Pro Lys Arg Val Ser  
50 55 60  
Arg Ile Ala Asn Arg Lys Ser Lys Lys Leu Ile Val Asn Glu Asp  
65 70 75 80  
Val Ala Gly Asn Tyr Asp Asp Thr Phe Gly Asp Val Gln Lys Gln Ile  
85 90 95  
Val Asn Tyr Phe Thr Tyr Lys Ala Val Arg Thr Val Leu His Gln Leu  
100 105 110  
Tyr Glu Met Asn Pro Pro Gln Tyr Thr Trp Phe Tyr Asn His Ile Ile  
115 120 125  
Thr Asn Arg Pro Thr Asp Gly Lys Arg Phe Leu Arg Ala Leu Gly Lys  
130 135 140  
Glu Ser Gln Glu Leu Ala Glu Arg Val Met Ile Thr Arg Leu His Leu  
145 150 155 160  
Tyr Gly Lys Trp Ile Lys Lys Cys Asp His Gly Lys Ile Tyr Gln Glu  
165 170 175  
Ile Ser Asp Glu Asn Leu Ala Leu Met Arg Glu Arg Leu Met Glu Thr  
180 185 190  
Val Ile Trp Pro Ser Asp Thr Asn Ser Arg  
195 200

&lt;210&gt; 214

&lt;211&gt; 191

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..191:Ceres Seq. ID 1396805

<400>214  
Met Asp Ser Ser Ser Pro Cys Leu Cys Leu Asp Ala His Thr Thr  
1 5 10 15  
Gly Thr Ile Arg Arg Lys Lys Ile Leu Gly Lys Ala Arg Asn Leu Glu  
20 25 30  
Leu Gly Ser Ser Phe Thr Gly Ser Arg Ile Val Phe Arg Leu Ser Pro  
35 40 45  
Lys Arg Val Ser Arg Ile Ala Asn Arg Lys Ser Lys Leu Leu Ile  
50 55 60  
Val Asn Glu Asp Val Ala Gly Asn Tyr Asp Asp Thr Phe Gly Asp Val  
65 70 75 80  
Gln Lys Gln Ile Val Asn Tyr Phe Thr Tyr Lys Ala Val Arg Thr Val  
85 90 95  
Leu His Gln Leu Tyr Glu Met Asn Pro Pro Gln Tyr Thr Trp Phe Tyr  
100 105 110  
Asn His Ile Ile Thr Asn Arg Pro Thr Asp Gly Lys Arg Phe Leu Arg  
115 120 125  
Ala Leu Gly Lys Glu Ser Gln Glu Leu Ala Glu Arg Val Met Ile Thr  
130 135 140  
Arg Leu His Leu Tyr Gly Lys Trp Ile Lys Lys Cys Asp His Gly Lys  
145 150 155  
Ile Tyr Gln Glu Ile Ser Asp Glu Asn Leu Ala Leu Met Arg Glu Arg  
165 170 175

101

SUBSTITUTE SHEET (RULE 26)

Leu Met Glu Thr Val Ile Trp Pro Ser Asp Thr Asn Ser Arg  
180 185 190

&lt;210&gt; 215

&lt;211&gt; 753

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..753:Ceres Seq. ID 1397130

<400>215  
aaagatttgg agaaaaagtt tactaaaaa gaatctgggg aagtaagtga gatgaagaa  
60  
ggggaggaga tgggaaacc agcgaacca aagccaaga agatgtttgc tctggggg  
120  
ttgattgata cctatgtcgc acagtgcgat agtgcaca agtgagggtt gattgatagc  
180  
caggaggaat atgaagatat cagaagtana atgtctcagg atccttttaa cgtcagaag  
240  
aaacaggaca tctcttctga agacctgtgt gatattgact acgattcttc tgggacttgg  
300  
gtcattgaca agcttggtct cccaacaac cctaaagttt tcaaggaag ctagttcttc  
360  
agaaaagatt actctaagat ggatacttac tactttact ctaccgggaa gaagctcagg  
420  
atgcgaatg aaatcgtcgc ctctgttgaa gccaatccgg aattcaggaa cgcaccatt  
480  
ggagacttca atttactgt cccaaggttc atggaagata ctgttcccc tgcaccgaag  
540  
ctgtgtcttc cttttccaag cactactacc actacttcag agaaagcag tctcaagcag  
600  
agccataact aatattgtct ctctccggg ccttttcta ctttcttct gttccctaa  
660  
gcaatgagtt ttgagctttt atagatttga tgttctctaa aactgttatg atattaaact  
720  
ctctaaagca aaaaatatga tcttttgtct ctc  
753

&lt;210&gt; 216

&lt;211&gt; 203

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..203:Ceres Seq. ID 1397131

<400>216  
Lys Asp Leu Glu Lys Lys Phe Thr Lys Lys Glu Ser Gly Glu Val Ser  
1 5 10 15  
Glu Met Lys Glu Glu Glu Ile Gly Lys Pro Ala Lys Pro Lys Ala  
20 25 30  
Lys Lys Asp Val Ala Pro Gly Arg Leu Ile Asp Thr Tyr Ala Ala Gln  
35 40 45  
Cys Asp Asn Cys His Lys Trp Arg Val Ile Asp Ser Gln Glu Tyr  
50 55 60  
Glu Asp Ile Arg Ser Lys Met Leu Glu Asp Pro Phe Asn Cys Gln Lys  
65 70 75 80  
Lys Gln Gly Met Ser Cys Glu Glu Pro Ala Asp Ile Asp Tyr Asp Ser  
85 90 95  
Ser Arg Thr Trp Val Ile Asp Lys Pro Gly Leu Pro Lys Thr Pro Lys  
100 105 110  
Gly Phe Lys Arg Ser Leu Val Leu Arg Lys Asp Tyr Ser Lys Met Asp  
115 120 125  
Thr Tyr Tyr Phe Thr Pro Thr Gly Lys Lys Leu Arg Ser Arg Asn Glu  
130 135 140  
Ile Ala Ala Phe Val Glu Ala Asn Pro Glu Phe Arg Asn Ala Pro Leu  
145 150 155  
Gly Asp Phe Asn Phe Thr Val Pro Lys Val Met Glu Asp Thr Val Pro  
165 170 175

102

SUBSTITUTE SHEET (RULE 26)

Pro Asp Pro Lys Leu Gly Ser Pro Phe Pro Ser Thr Thr Thr Thr Thr  
180 185 190  
Ser Glu Lys Ser Ser Val Lys Gln Ser His Asn  
195 200

&lt;210&gt; 217

&lt;211&gt; 186

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..186:Ceres Seq. ID 1397132

&lt;400&gt;217

Met Lys Glu Glu Glu Glu Ile Gly Lys Pro Ala Lys Pro Lys Ala Lys  
1 5 10 15

Lys Asp Val Ala Pro Gly Arg Leu Ile Asp Thr Tyr Ala Ala Gln Cys

Asp Asn Cys His Lys Trp Arg Val Ile Asp Ser Gln Glu Glu Tyr Glu  
35 40 45

Asp Ile Arg Ser Lys Met Leu Glu Asp Pro Phe Asn Cys Gln Lys Lys  
50 55 60

Gln Gly Met Ser Cys Glu Glu Pro Ala Asp Ile Asp Tyr Asp Ser Ser  
65 70 75 80

Arg Thr Trp Val Ile Asp Lys Pro Gly Leu Pro Lys Thr Pro Lys Gly  
85 90 95

Phe Lys Arg Ser Leu Val Leu Arg Lys Asp Tyr Ser Lys Met Asp Thr  
100 105 110

Tyr Tyr Phe Thr Pro Thr Gly Lys Lys Leu Arg Ser Arg Asn Glu Ile  
115 120 125

Ala Ala Phe Val Glu Ala Asn Pro Glu Phe Arg Asn Ala Pro Leu Gly  
130 135 140

Asp Phe Asn Phe Thr Val Pro Lys Val Met Glu Asp Thr Val Pro Pro  
145 150 155 160

Asp Pro Lys Leu Gly Ser Pro Phe Pro Ser Thr Thr Thr Thr Ser  
165 170 175

Glu Lys Ser Ser Val Lys Gln Ser His Asn  
180 185

&lt;210&gt; 218

&lt;211&gt; 133

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..133:Ceres Seq. ID 1397133

&lt;400&gt;218

Met Leu Glu Asp Pro Phe Asn Cys Gln Lys Lys Gln Gly Met Ser Cys  
1 5 10 15

Glu Glu Pro Ala Asp Ile Asp Tyr Asp Ser Ser Arg Thr Trp Val Ile  
20 25 30

Asp Lys Pro Gly Leu Pro Lys Thr Pro Lys Gly Phe Lys Arg Ser Leu  
35 40 45

Val Leu Arg Lys Asp Tyr Ser Lys Met Asp Thr Tyr Thr Phe Thr Pro  
50 55 60

Thr Gly Lys Lys Leu Arg Ser Arg Asn Glu Ile Ala Ala Phe Val Glu

65 Ala Asn Pro Glu Phe Arg Asn Ala Pro Leu Gly Asp Phe Asn Phe Thr  
70 75 80  
Val Pro Lys Val Met Glu Asp Thr Val Pro Pro Asp Pro Lys Leu Gly  
100 105 110

Ser Pro Phe Pro Ser Thr Thr Thr Thr Ser Glu Lys Ser Ser Val  
115 120 125

Lys Gln Ser His Asn  
130

&lt;210&gt; 219

&lt;211&gt; 784

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..784:Ceres Seq. ID 1398004

&lt;400&gt;219

actcatcaac caaacacaaa cataaaaaa caagtgaag ctttaaaacg agaggagag  
60  
agcaaaatg gcgaatcgg gaactacgt gacgaagtt ccgtataaag gatcgccga  
120  
gaatactac aagagtgga agaagcgaa ccatgttcc cctgatcta tcggccacca  
180  
catccaatg gttacgttc acgaagcga acatgactct cacgggtcta tcaggagtgc  
240  
gaactacaa tgggatgaa aggaagagt gtccaagag agaagagaga tagacagtga  
300  
gacaaaacg ttgacttaa ggggacttga gggtaactgc atgggacgc tcaagtgtta  
360  
cgactgcgc taccattca ttccaatc ttgagatcc tgcatacga aatcaacttt  
420  
aatatggag aagcgaacg atgattccc agaaccaagc ggtacataa aattcgtcaa  
480  
gagcttggt gctgacatgg gaacaacgt tagcaaaact taatcatcat tcccacagtc  
540  
gtcgtcgtc tcatcatcat tcatcatc atcatcatca tcatcatcat catcatc  
600  
atcatcata ctatctgat tgaatgta agatgtttc agtataata atgggtctt  
660  
gtgagtcgt cattctatg tgaacacgt ttggtctgt atgatcttc gatataatg  
720  
tatgtcatg atcatatgct ggttcgata taatgattct taagtataat ttactacaa  
780  
tttc 784

&lt;210&gt; 220

&lt;211&gt; 151

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..151:Ceres Seq. ID 1398005

&lt;400&gt;220

Met Ala Thr Ser Gly Thr Tyr Val Thr Glu Val Pro Leu Lys Gly Ser  
1 5 10 15

Ala Glu Lys Tyr Tyr Lys Arg Trp Lys Asn Glu Asn His Val Phe Pro  
20 25 30

Asp Ala Ile Gly His His Ile Gln Asn Val Thr Val His Glu Gly Glu  
35 40 45

His Asp Ser His Gly Ser Ile Arg Ser Trp Asn Tyr Thr Trp Asp Gly  
50 55 60

Lys Glu Glu Val Phe Lys Glu Arg Arg Glu Ile Asp Asp Glu Thr Lys  
65 70 75 80

Thr Leu Thr Leu Arg Gly Leu Glu Gly His Val Met Glu Gln Leu Lys  
85 90 95

Val Tyr Asp Val Val Tyr Gln Phe Ile Pro Lys Ser Glu Asp Thr Cys  
100 105 110

Ile Gly Lys Ile Thr Leu Ile Trp Glu Lys Arg Asp Ser Pro  
115 120  
Glu Pro Ser Gly Tyr Ile Lys Phe Val Lys Ser Leu Val Ala Asp Met  
130 140  
Gly Asn His Val Ser Lys Thr  
145 150

&lt;210&gt; 221

&lt;211&gt; 389

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..389:Ceres Seq. ID 1399370

&lt;400&gt;221

aaaaatttt tcagagtcgg aggtttcaga cagagagaa gaagatgaag cggacggtag 60  
tggacaacg gattgatca tcgtgggtgg tgttgatc attggccttt ggtacttgt 120  
cactggagct tggttacaag cctttcttg aaaggtga acaatacga agatctcttc 180  
agttctca acaatcaaa caacaagatg acaagaaga agcaggtgg gacaaagca 240  
atgcagggg gtggggggg aaaggttgt cttataatg gttcaatc agaatatca 300  
aagtgtaga ggaagaaga ccttatcaaa attttctct taatcaacg taagtcaat 360  
ctgttaacta tcaaccatgg cctgtttt 389

&lt;210&gt; 222

&lt;211&gt; 88

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..88:Ceres Seq. ID 1399371

&lt;400&gt;222

Lys Leu Phe Gln Ser Arg Arg Phe Gln Thr Glu Arg Lys Lys Met Lys  
1 5 10 15  
Arg Thr Val Met Asp Asn Ala Ile Arg Ser Ser Val Val Leu Gly

Ser Leu Ala Phe Gly Tyr Leu Ser Leu Glu Leu Gly Tyr Lys Pro Phe  
35 40 45

Leu Glu Lys Ala Glu Gln Tyr Glu Arg Ser Leu Gln Ser Gln Gln  
50 55 60

His Gln Gln Gln Asp Glu Gln Glu Ala Arg Trp Asp Asn Ser Asn  
65 70 75 80

Val Glu Gly Trp Glu Glu Lys Arg  
85

&lt;210&gt; 223

&lt;211&gt; 74

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..74:Ceres Seq. ID 1399372

&lt;400&gt;223

Met Lys Arg Thr Val Met Asp Asn Ala Ile Arg Ser Ser Val Val Val  
1 5 10 15  
Leu Gly Ser Leu Ala Phe Gly Tyr Leu Ser Leu Glu Gly Tyr Lys  
20 25 30  
Pro Phe Leu Glu Lys Ala Glu Gln Tyr Glu Arg Ser Leu Gln Ser Ser  
35 40 45  
Gln Gln His Gln Gln Asp Glu Gln Glu Ala Arg Trp Asp Asn  
50 55 60  
Ser Asn Val Glu Gly Trp Glu Glu Lys Arg  
65 70

&lt;210&gt; 224

&lt;211&gt; 69

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..69:Ceres Seq. ID 1399373

&lt;400&gt;224

Met Asp Asn Ala Ile Arg Ser Ser Val Val Val Leu Gly Ser Leu Ala  
1 5 10 15  
Phe Gly Tyr Leu Ser Leu Glu Leu Gly Tyr Lys Pro Phe Leu Glu Lys  
20 25 30  
Ala Glu Gln Tyr Glu Arg Ser Leu Gln Ser Ser Gln Gln His Gln Gln  
35 40 45  
Gln Asp Glu Gln Glu Ala Arg Trp Asp Asn Ser Asn Val Glu Gly  
50 55 60  
Trp Glu Glu Lys Arg  
65

&lt;210&gt; 225

&lt;211&gt; 745

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..745:Ceres Seq. ID 1425147

&lt;400&gt;225

attcttcgt ctcgtctcc aaatogaat caaatctct aaagtcca ttttttgtt 60  
ctgtctttt ttttttta aagaatggt tcaattctg caacttgc ttgcctattg 120  
ttaactcac agagaaaatc caatccaca tcgattcaaa aactccatt tctctaact 180  
caggtaaga atgatcttc tccattatct ctactcga accctagcag cctagctcg 240  
atgcagaat ctadtgaga aagtcagat tcatcgact ttatcgact tgttagtgc 300  
attcagaatg ttgggataa gtcgaagat aggttagtc ttattggtt ggttttgt 360  
ggtattgtag cttcttggc atcatgaat ctatcaagg caattgcaa attgccegt 420  
atctcagcg ctttcgaact agttgtatc ttgttcca cgtgttcac atatcgatat 480  
ctctgttca aaccggacg acaggactt tcgaaatg tcaagaatc agtagcgat 540  
atactggcc agtaacctt gtgtgtga taacttca tcttggag agtatttgt 600  
tgcaagttg taatatca tgaacggggt gtgtgtgtt ctgctccat aatgcacg 660  
attgaaacc tgtaataact ttattgtg ttttgggtg tgagaanaa caatcttcc 720  
taattcaaa gattctctt tatgt 745

&lt;210&gt; 226

&lt;211&gt; 156

&lt;212&gt; PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..156:Ceres Seq. ID 1425148

<400>226

Met Ala Ser Ile Ser Ala Thr Leu Pro Ser Pro Leu Leu Thr Gln

1 Arg Lys Ser Asn Leu Thr Ser Ile Gln Lys Leu Pro Phe Ser Leu Thr

20 25 30

Arg Gly Thr Asn Asp Leu Ser Pro Leu Ser Leu Thr Arg Asn Pro Ser

35 40 45

Ser Ile Ser Leu Met Val Lys Ser Ser Gly Glu Ser Ser Asp Ser Ser

50 55 60

Thr Asp Leu Asp Val Val Ser Thr Ile Gln Asn Val Trp Asp Lys Ser

65 70 80

Glu Asp Arg Leu Gly Leu Ile Gly Leu Gly Phe Ala Gly Ile Val Ala

85 90 95

Leu Trp Ala Ser Leu Asn Leu Ile Thr Ala Ile Asp Lys Leu Pro Val

100 105 110

Ile Ser Ser Gly Phe Glu Leu Val Gly Ile Leu Phe Ser Thr Trp Phe

115 120 125

Thr Tyr Arg Tyr Leu Leu Phe Lys Pro Asp Arg Gln Glu Leu Ser Lys

130 135 140

Ile Val Lys Lys Ser Val Ala Asp Ile Leu Gly Gln

145 150 155

<210> 227

<211> 104

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..104:Ceres Seq. ID 1425149

<400>227

Met Val Lys Ser Ser Gly Glu Ser Ser Asp Ser Ser Thr Asp Leu Asp

1 Val Val Ser Thr Ile Gln Asn Val Trp Asp Lys Ser Glu Asp Arg Leu

20 25 30

Gly Leu Ile Gly Leu Gly Phe Ala Gly Ile Val Ala Leu Trp Ala Ser

35 40 45

Leu Asn Leu Ile Thr Ala Ile Asp Lys Leu Pro Val Ile Ser Ser Gly

50 55 60

Phe Glu Leu Val Gly Ile Leu Phe Ser Thr Trp Phe Thr Tyr Arg Tyr

65 70 75 80

Leu Leu Phe Lys Pro Asp Arg Gln Glu Leu Ser Lys Ile Val Lys Lys

85 90 95

Ser Val Ala Asp Ile Leu Gly Gln

100

<210> 228

<211> 769

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..769:Ceres Seq. ID 1441102

<400>228

ccacacaca acttcagctt taagacagct aaaaactaat acaagaacac agagatgggt

1 ttggtacag agaggtgag agctaaggca gagatgtaca ccggagatga gatatgtaga

20 25 30

gagagacaa agtgttctt caagaaata ttatgtacca atgtttatt accattgaag

240 gatagaag aggttggtta tgacagagag taagtggttg tatgtctgaa gcagaagaag

280 agatacatc acaattcac agagatagat aagttgtct cctatgaac cgaagtcaaa

320 gcaatgttg agaccggaa aatcaagaag ctcaactggag tcaagccaa ggaacttctt

360 atttggtca ctatcaatga gatcatacc gaggaacccc ctaccaagt cagcttcaag

400 acacagacca cactgtccag gactttccc gtccacagctt ttatagtccc agaagaacct

440 gccaggagg acctcgcaa agaggacct gccaggaga agagcagcga agccaccgag

480 gccaggagg ctgttgcaat caagagctg gtccagctca aagagcgcc ctactcagct

520 cagcttcagc agataaaga gaccgaatt tttaagtgtt ctcaagaat ttgaacaaat

560 caggagaggt tagtgttctt cactgtctta actatcatgt attatcaccc taccacttgt

600 tgtgtgttgt tttaactgtta aatagatc aatacatcc tgttccact

640

<210> 229

<211> 197

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..197:Ceres Seq. ID 1441103

<400>229

Pro Gln His Asn Phe Ser Phe Lys Thr Ala Lys Thr Lys Tyr Lys Glu

1 Thr Glu Met Gly Leu Val Thr Glu Glu Val Arg Ala Lys Ala Glu Met

20 25 30

Tyr Thr Gly Asp Glu Ile Cys Arg Glu Lys Thr Lys Cys Phe Leu Lys

35 40 45

Glu Ile Ser Met Pro Asn Gly Leu Leu Pro Leu Lys Asp Ile Glu Glu

50 55 60

Val Gly Tyr Asp Arg Glu Ser Gly Val Val Trp Leu Lys Lys Lys

65 70 75 80

Ser Ile Thr His Lys Phe Thr Glu Ile Asp Lys Leu Val Ser Tyr Gly

85 90 95

Thr Glu Val Thr Ala Ile Val Glu Thr Gly Lys Ile Lys Lys Leu Thr

100 105 110

Gly Val Lys Ala Lys Glu Leu Leu Ile Trp Val Thr Ile Asn Glu Ile

115 120 125

Tyr Thr Glu Glu Pro Thr Lys Ile Thr Phe Lys Thr Pro Thr Thr

130 135 140

Leu Ser Arg Thr Phe Pro Val Thr Ala Phe Ile Val Pro Glu Glu Pro

145 150 155 160

Ala Lys Glu Glu Pro Ala Lys Glu Glu Pro Ala Lys Glu Lys Ser Ser

165 170 175

Glu Ala Thr Glu Ala Lys Glu Ala Val Ala Ile Lys Glu Ala Val Ala

180 185 190

Val Lys Glu Ala Ala

195

<210> 230

<211> 179

<212> PRT

<213> Arabidopsis thaliana



<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..179:Ceres Seq. ID 1441104

<400>230  
 Met Gly Leu Val Thr Glu Glu Val Arg Ala Lys Ala Glu Met Tyr Thr  
 1 10 15  
 Gly Asp Glu Ile Cys Arg Glu Lys Thr Lys Cys Phe Leu Lys Glu Ile  
 20 25 30  
 Ser Met Pro Asn Gly Leu Leu Pro Leu Lys Asp Ile Glu Glu Val Gly  
 35 40 45  
 Tyr Asp Arg Glu Ser Gly Val Val Trp Leu Lys Glu Lys Lys Ser Ile  
 50 55 60  
 Thr His Lys Phe Thr Glu Ile Asp Lys Leu Val Ser Tyr Gly Thr Glu  
 65 70 75 80  
 Val Thr Ala Ile Val Glu Thr Gly Lys Ile Lys Lys Leu Thr Gly Val  
 85 90 95  
 Lys Ala Lys Glu Leu Ile Trp Val Thr Ile Asn Glu Ile Tyr Thr  
 100 105 110  
 Glu Glu Pro Pro Thr Lys Ile Thr Phe Lys Thr Pro Thr Thr Leu Ser  
 115 120 125  
 Arg Thr Phe Pro Val Thr Ala Phe Ile Val Pro Glu Glu Pro Ala Lys  
 130 135 140  
 Glu Glu Pro Ala Lys Glu Glu Pro Ala Lys Glu Lys Ser Ser Glu Ala  
 145 150 155 160  
 Thr Glu Ala Lys Glu Ala Val Ala Ile Lys Glu Ala Val Ala Val Lys  
 165 170 175  
 Glu Ala Ala

<210> 231

<211> 166

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..166:Ceres Seq. ID 1441105

<400>231  
 Met Tyr Thr Gly Asp Glu Ile Cys Arg Glu Lys Thr Lys Cys Phe Leu  
 1 5 10 15  
 Lys Glu Ile Ser Met Pro Asn Gly Leu Leu Pro Leu Lys Asp Ile Glu  
 20 25 30  
 Glu Val Gly Tyr Asp Arg Glu Ser Gly Val Val Trp Leu Lys Glu Lys  
 35 40 45  
 Lys Ser Ile Thr His Lys Phe Thr Glu Ile Asp Lys Leu Val Ser Tyr  
 50 55 60  
 Gly Thr Glu Val Thr Ala Ile Val Glu Thr Gly Lys Ile Lys Lys Leu  
 65 70 75 80  
 Thr Gly Val Lys Ala Lys Glu Leu Ile Trp Val Thr Ile Asn Glu  
 85 90 95  
 Ile Tyr Thr Glu Glu Pro Thr Lys Ile Thr Phe Lys Thr Pro Thr  
 100 105 110  
 Thr Leu Ser Arg Thr Phe Pro Val Thr Ala Phe Ile Val Pro Glu Glu  
 115 120 125  
 Pro Ala Lys Glu Glu Pro Ala Lys Glu Glu Pro Ala Lys Glu Lys Ser  
 130 135 140

Ser Glu Ala Thr Glu Ala Lys Glu Ala Val Ala Ile Lys Glu Ala Val  
 145 150 155 160  
 Ala Val Lys Glu Ala Ala  
 165

<210> 232

<211> 607

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..607:Ceres Seq. ID 1447480

<400>232  
 aaatctcta aagtttcaat tttttgttc tttttttt tttttaag aatggctta  
 attctgcaa atttgcttc gcaatgta ctcaacaga gaaatccaa tctcacatcg  
 attcaaac tccattttc tctaactga gtaacgaatg attttctcc attatctctt  
 actgaacc ctgacgat cagctcgatg gtgaagcta gtggagaag ctgacttca  
 tgcctgac tcgacgttgt tagtacgatg cagaatgtg caatigacaa attgccgtt  
 actcagcag gattcgaaact agttgtatc ttgtctca cgtgttcac atatcgatat  
 ctctgttca accggacag acaggagctt tcgaaatgy tcaagaatc agtagcgat  
 atacttgcc agtgaacct gtgtgtgtga taacttca ttttggag atgattgtt  
 tgaagttg taataatca tgaacgggtg gtgtgtgtt ctagtccaat aatgcatgc  
 attgaacc tgaataact ttatgtgtg ttttgtgtg tgagcaaat caatctttc  
 taatttc

<210> 233

<211> 127

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..127:Ceres Seq. ID 1447481

<400>233  
 Met Ala Ser Ile Ser Ala Thr Leu Pro Ser Pro Leu Leu Thr Gln  
 1 5 10 15  
 Arg Lys Ser Asn Leu Thr Ser Ile Gln Lys Leu Pro Phe Ser Leu Thr  
 20 25 30  
 Arg Gly Thr Asn Asp Leu Ser Pro Leu Ser Leu Thr Arg Asn Pro Ser  
 35 40 45  
 Ser Ile Ser Leu Met Val Lys Ala Ser Gly Glu Ser Ser Asp Ser Ser  
 50 55 60  
 Thr Asp Leu Asp Val Val Ser Thr Ile Gln Asn Val Ala Ile Asp Lys  
 65 70 75 80  
 Leu Pro Val Ile Ser Ser Gly Phe Glu Leu Val Gly Ile Leu Phe Ser  
 85 90 95  
 Thr Trp Phe Thr Tyr Arg Tyr Leu Phe Lys Pro Asp Arg Gln Glu  
 100 105 110  
 Leu Ser Lys Ile Val Lys Lys Ser Val Ala Asp Ile Leu Gly Gln  
 115 120 125

<210> 234

<211> 75

<212> PRT

<213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..75:Ceres Seq. ID 1447482

<400>234

Met Val Lys Ala Ser Gly Glu Ser Ser Asp Ser Ser Thr Asp Leu Asp  
 1 5 10 15  
 Val Val Ser Thr Ile Gln Asn Val Ala Ile Asp Lys Leu Pro Val Ile  
 20 25 30  
 Ser Ser Gly Phe Glu Leu Val Gly Ile Leu Phe Ser Thr Thr Phe Thr  
 35 40 45  
 Tyr Arg Tyr Leu Leu Phe Lys Pro Asp Arg Gln Glu Leu Ser Lys Ile  
 50 55 60  
 Val Lys Lys Ser Val Ala Asp Ile Leu Gly Gln  
 65 70 75

<210> 235

<211> 668

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..668:Ceres Seq. ID 1447577

<400>235

aaaaaaac aaaaacaaa attattattc agaaaaaag gaaaaatga atttcatctc  
 60  
 cgtacaggta aagaactct caagctcaac acagaggcc agaccacaac aagccagtcg  
 120  
 aggaaccca aacagtatac agccagcta ccaagccga gctcatggca agtgcacaag  
 180  
 ttgatgta agctgtctca gccgcagtc gtaacgaatc agacaactc gcaagggtg  
 240  
 aagtcgcgg agctcttgt gatatcttag acgtgccga gaatacgggt aagttcgatg  
 300  
 aaagagtac cactgtgtag taactgaca aggtcgagaa gatactaac gactacaggt  
 360  
 cgtcacact caccgtgtct ggfggctct cctctccac gagtcaaggt gaggcagcaa  
 420  
 gtcacctga gccggcggtc aagaagagc atgaagatc tgggtgtggg cttggagtt  
 480  
 atgcagaat ggtccaagt ttctgaagt gatctgct ttaattgtg ttcatcatt  
 540  
 tcgtaataa aaataaata actagatcgt ttgtgacta gtttatgtg cttggtttat  
 600  
 gttatgggg agtgacgagt gadtgaata actctgtg atcatgaatc taatccatct  
 660  
 ttgtgtgc

<210> 236

<211> 62

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..62:Ceres Seq. ID 1447578

<400>236

Lys Lys Asn Lys Thr Lys Asn Tyr Ile Gln Glu Lys Lys Glu Lys Met  
 1 5 10 15  
 Asn Phe Ile Ser Asp Gln Val Lys Lys Leu Ser Ser Ser Thr Pro Glu  
 20 25 30  
 Ala Arg Pro Gln Gln Ala Ser Arg Arg Asn Arg Asn Ser Tyr Lys Thr  
 35 40 45  
 Ser Tyr Gln Arg Arg Ala His Gly Lys Cys Gln Gly Cys Ser  
 50 55 60

<210> 237  
 <211> 115  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..115:Ceres Seq. ID 1447579

<400>237

Met Ala Ser Ala Lys Val Val Ala Glu Ala Ala Gln Ala Ala Ala Arg  
 1 5 10 15  
 Asn Glu Ser Asp Lys Leu Asp Lys Gly Lys Val Ala Gly Ala Ser Ala  
 20 25 30  
 Asp Ile Leu Asp Ala Ala Glu Lys Tyr Gly Lys Phe Asp Glu Lys Ser  
 35 40 45  
 Ser Thr Gly Gln Tyr Leu Asp Lys Ala Glu Lys Tyr Leu Asn Asp Tyr  
 50 55 60  
 Glu Ser Ser His Ser Thr Gly Ala Gly Gly Pro Pro Pro Thr Ser  
 65 70 75  
 Gln Ala Glu Pro Ala Ser Gln Pro Glu Pro Ala Ala Lys Lys Asp Asp  
 85 90 95  
 Glu Glu Ser Gly Gly Leu Gly Tyr Ala Lys Met Ala Gln Gly  
 100 105 110  
 Phe Leu Lys  
 115

<210> 238

<211> 69

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..69:Ceres Seq. ID 1447580

<400>238

Met Lys Arg Val Ala Leu Val Ser Thr Ser Thr Arg Leu Arg Ser Ile  
 1 5 10 15  
 Ser Thr Thr Thr Ser Arg His Thr Pro Pro Val Leu Val Val Leu Leu  
 20 25 30  
 Leu Arg Arg Val Arg Leu Ser Gln Gln Val Ser Leu Ser Arg Arg Leu  
 35 40 45  
 Arg Lys Thr Met Lys Ser Leu Val Val Gly Leu Glu Val Met Pro Arg  
 50 55 60  
 Trp Leu Lys Val Ser  
 65

<210> 239

<211> 717

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..717:Ceres Seq. ID 1447922

<400>239

cttgctctct gtctcaaaa tcgaatacaa atctctaag tttcaatttt ttgttctgt 60  
 tctttttt ttftaaaga atggtttcaa ttctgcaac ttgtctcttg ccaattgtac 120  
 tcacacag aaatacaat ctacatacga ttcaaaact ccaattttct ctactcag 180  
 gtacacaga tctttctca ttatctcta ctgcaaac ccagcagcag agtctatgg 240  
 tgaactag tgaagaagc tcaatttat caactatct caactgtgt agtactatc 300  
 agaatgttg gataagctc gaagataggt taggtcttat ttgtttggt ttgtcgtga 360  
 ttgtactct ttggcatca ttgaatcga tcaaggcaat tgacaattg ccggttatct 420  
 cgagcgatt cgaactagtt ggtacttgt tctcaactgt gttcaactat cgatctct 480  
 tttcaaac ggcagacag gacttttga aaattgtcaa gaatacagta gcygatatac 540  
 ttggcagtg aacttgtgt gttgataat acttcatct ttgaagatga ttgtttgca 600  
 agtttcaa attacatgac aggttggttg ttgtttctag tccaataag tcatgcat 660  
 gaactcga aatacttat tgtttgtttt tggttgtgag caaatcaat cttttct 717

&lt;210&gt; 240

&lt;211&gt; 156

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..156:Ceres Seq. ID 1447923

&lt;400&gt;240

Met Ala Ser Ile Ser Ala Thr Leu Pro Ser Pro Leu Leu Leu Thr Gln 15  
 1 Arg Lys Ser Asn Leu Thr Ser Ile Gln Lys Leu Pro Phe Ser Leu Thr 30  
 20 Arg Gly Thr Asn Asp Leu Ser Pro Leu Ser Leu Thr Arg Asn Pro Ser 45  
 35 Ser Ile Ser Leu Met Val Lys Ala Ser Gly Gln Ser Ser Asp Ser Ser 60  
 50 Thr Asp Leu Asp Val Val Ser Thr Ile Gln Asn Val Trp Asp Lys Ser 80  
 65 Glu Asp Arg Leu Gly Leu Ile Gly Leu Gly Phe Ala Gly Ile Val Ala 90  
 85 Leu Trp Ala Ser Leu Asn Leu Ile Thr Ala Ile Asp Lys Leu Pro Val 110  
 100 Ile Ser Ser Gly Phe Glu Leu Val Gly Ile Leu Phe Ser Thr Trp Phe 125  
 115 Thr Tyr Arg Tyr Leu Leu Phe Lys Pro Asp Arg Gln Glu Leu Ser Lys 140  
 130 Ile Val Lys Lys Ser Val Ala Asp Ile Leu Gly Gln 155  
 145

&lt;210&gt; 241

&lt;211&gt; 104

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..104:Ceres Seq. ID 1447924

&lt;400&gt;241

Met Val Lys Ala Ser Gly Glu Ser Ser Asp Ser Ser Thr Asp Leu Asp 15  
 1 Val Val Ser Thr Ile Gln Asn Val Trp Asp Lys Ser Glu Asp Arg Leu 30  
 20 Gly Leu Ile Gly Leu Gly Phe Ala Gly Ile Val Ala Leu Trp Ala Ser 45

35 Leu Asn Leu Ile Thr Ala Ile Asp Lys Leu Pro Val Ile Ser Ser Gly 45  
 50 Phe Glu Leu Val Gly Ile Leu Phe Ser Thr Trp Phe Thr Tyr Arg Tyr 60  
 65 Leu Leu Phe Lys Pro Asp Arg Gln Glu Leu Ser Lys Ile Val Lys Lys 80  
 85 Ser Val Ala Asp Ile Leu Gly Gln 100

&lt;210&gt; 242

&lt;211&gt; 656

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..656:Ceres Seq. ID 1448012

&lt;400&gt;242

ggaattctc actctctcta tctctcttag ccagctctga ccatttggat tttttccgg 60  
 tgaagaaggga gcagaaacat ggtgtgtact atcagattat ccagatttgg atgcaaaat 120  
 cggcatttt ttagggttat ggtgtgtact agcagatctc caagagacgg gaagcattt 180  
 gaggtcttag gttacttcaa tctttgccca ggcacggagc gtgtaagag gatgggtctc 240  
 aggtctatc gaattaagta ctgttatct gtgtgtctc agccatcaga ccggtttcaa 300  
 cgcctcttt tcaatcggc tttacttct cctctccaa tgggtgctat gggacgtaaa 360  
 ggtgagcac gagacacacg cccatttgt ccaatgactg tctgtatgt ggaatgcagag 420  
 aataaacag ttaatgccaa tgaataccag cctaaggag aggtatcaga agacaagatt 480  
 gaatgatca ttactctct gtctgttag cttttcaagt tcaatttct gcgatata 540  
 ttgttaatg cagcataga caactgact ttctcttgg ttggcgata aacggcaagg 600  
 tgttggcnc tttttgaga aacggcaact atttgcatt gggatatttt aatttt 656

&lt;210&gt; 243

&lt;211&gt; 135

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..135:Ceres Seq. ID 1448013

&lt;400&gt;243

Met Val Val Arg Ile Arg Leu Ser Arg Phe Gly Cys Lys Asn Arg Pro 15  
 1 Phe Phe Arg Val Met Ala Ala Asp Ser Arg Ser Pro Arg Asp Gly Lys 30  
 20 His Leu Glu Val Leu Gly Tyr Phe Asn Pro Leu Pro Gly Gln Asp Gly 45  
 35 Gly Lys Arg Met Gly Leu Lys Phe Asp Arg Ile Lys Tyr Trp Leu Ser 60  
 50 Val Gly Ala Gln Pro Ser Asp Pro Val Gln Arg Leu Leu Phe Arg Ser 75  
 65 Gly Leu Leu Pro Pro Pro Met Val Ala Met Gly Arg Lys Gly Gly 80  
 85 Ala Arg Asp Thr Arg Pro Val Asp Pro Met Thr Gly Arg Tyr Val Asp 95  
 100 Ala Glu Asn Lys Thr Val Asn Ala Asn Asp Asn Gln Pro Lys Glu Glu 110  
 115 Asp Ser Glu Asp Lys Ile Ala 120  
 125

130 135

<210> 244  
 <211> 115  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..115:Ceres Seq. ID 1448014

<400>244  
 Met Ala Ala Asp Ser Arg Ser Pro Arg Asp Gly Lys His Leu Glu Val  
 1 5 10 15  
 Leu Gly Tyr Phe Asn Pro Leu Pro Gly Gln Asp Gly Lys Arg Met  
 20 25 30  
 Gly Leu Lys Phe Asp Arg Ile Lys Tyr Trp Leu Ser Val Gly Ala Gln  
 35 40 45  
 Pro Ser Asp Pro Val Gln Arg Leu Leu Phe Arg Ser Gly Leu Leu Pro  
 50 55 60  
 Pro Pro Pro Met Val Ala Met Gly Arg Lys Gly Gly Ala Arg Asp Thr  
 65 70 75  
 Arg Pro Val Asp Pro Met Thr Gly Arg Tyr Val Asp Ala Gln Asn Lys  
 85 90 95  
 Thr Val Asn Ala Asn Asp Asn Gln Pro Lys Glu Glu Asp Ser Glu Asp  
 100 105 110  
 Lys Ile Ala  
 115

<210> 245  
 <211> 84  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..84:Ceres Seq. ID 1448015

<400>245  
 Met Gly Leu Lys Phe Asp Arg Ile Lys Tyr Trp Leu Ser Val Gly Ala  
 1 5 10 15  
 Gln Pro Ser Asp Pro Val Gln Arg Leu Leu Phe Arg Ser Gly Leu Leu  
 20 25 30  
 Pro Pro Pro Pro Met Val Ala Met Gly Arg Lys Gly Gly Ala Arg Asp  
 35 40 45  
 Thr Arg Pro Val Asp Pro Met Thr Gly Arg Tyr Val Asp Ala Glu Asn  
 50 55 60  
 Lys Thr Val Asn Ala Asn Asp Asn Gln Pro Lys Glu Glu Asp Ser Glu  
 65 70 75  
 Asp Lys Ile Ala

<210> 246  
 <211> 458  
 <212> DNA  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..458:Ceres Seq. ID 1448135

<400>246  
 aaattctag ggttttgag agtgtctctc gtgcgcgttg taattctct gttagcaaat  
 60  
 cgacaaatg ggtcactcta atgtatgaa ctctcatccg aagaagtacg gtctcgatc  
 120  
 tegtattgc cgtgtgtcg ggaactgca cggctgac cgaagtatg gtttgaactg  
 180  
 ctgcagacac tgtttcogta gcaacgttaa ggaattgga ttcatlaagt accgttaate  
 240  
 aagcaccac ttcatgattg atgtctaatg atataaacat gaagcgttcg atgggttgg  
 300  
 ctttaagct ttgtagtt ttgaaattt tacttttgag aaccattgtt atttggag  
 360  
 ttaattaagt tgttgaact ctcattaagc atgtcttatt ttggattaat gatgtttgg  
 420  
 ctattctg attttgtt tctacgca atttgact

<210> 247  
 <211> 51  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..51:Ceres Seq. ID 1448136

<400>247  
 Phe Ser Arg Val Leu Lys Ser Val Ser Arg Arg Arg Cys Asn Ser Ser  
 1 5 10 15  
 Val Ser Lys Ser Thr Lys Trp Val Thr Leu Met Tyr Gly Thr Leu Ile  
 20 25 30  
 Arg Arg Ser Thr Val Leu Asp Leu Val Tyr Ala Val Cys Ala Gly Thr  
 35 40 45  
 Arg Thr Gly  
 50

<210> 248  
 <211> 56  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..56:Ceres Seq. ID 1448137

<400>248  
 Met Gly His Ser Asn Val Trp Asn Ser His Pro Lys Lys Tyr Gly Pro  
 1 5 10 15  
 Gly Ser Arg Leu Cys Arg Val Cys Gly Asn Ser His Gly Leu Ile Arg  
 20 25 30  
 Lys Tyr Gly Leu Asn Cys Cys Arg Gln Cys Phe Arg Ser Asn Ala Lys  
 35 40 45  
 Glu Ile Gly Phe Ile Lys Tyr Arg  
 50 55

<210> 249  
 <211> 38  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..38:Ceres Seq. ID 1448138

<400>249

Met Glu Leu Ser Ser Glu Glu Val Arg Ser Trp Ile Ser Phe Met Pro  
1 5 10 15  
Cys Val Arg Glu Leu Ala Arg Ala Asp Pro Glu Val Trp Phe Glu Leu  
20 25 30  
Leu Gln Thr Val Phe Pro  
35

<210> 250

<211> 603

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..603:Ceres Seq. ID 1448185

<400>250

gaagaagaa gtaetcttgt gagtttggg gatccgaaa atggaggttc caggttcac  
gaagaagatg atcgaagc agaaagagat gtcgcagct aaatagcac ttgattctag  
agatattgct gctactctct tgattccgct caacaatgt cgtcaggctg agtttaacct  
tccatgaa tgtaagagc agcgcacgt ttatgaagag ttgaaatagc agcttgat  
ggagaagatg ctgcgata gaagatccat tgaagaaga gctttgcta aacagaata  
actacaaga aagctgctg ttcctctct cctaaact gctaagctt aggtatcgat  
tctttcca attgcgct tccagattcc ggattctct ggaactgta agatgtgg  
ggctctgct ttcaactctt tctcttgc tagtgatga aaattgtac tacatttca  
gatgtgaag tcaagattc tctactctt tttttttt gaactcttt taactggag  
tggttccca aaaaataaga tgcanaactc atcttttgt tggtttcta tctttaact  
gtg

<210> 251

<211> 103

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..103:Ceres Seq. ID 1448186

<400>251

Met Glu Val Pro Gly Ser Ser Lys Lys Met Ile Ala Thr Gln Glu Glu  
1 5 10 15  
Met Ser Ala Ala Lys Ile Ala Leu Gly Ser Arg Asp Met Cys Ala His  
20 25 30  
Leu Leu Ile Pro Leu Asn Lys Cys Arg Gln Ala Glu Phe Tyr Leu Pro  
35 40 45  
Trp Lys Cys Glu Asp Glu Arg His Val Tyr Glu Lys Cys Glu Tyr Glu  
50 55 60  
Leu Val Met Glu Arg Met Leu Ala Met Lys Lys Ile Arg Glu Glu  
65 70 75 80  
Ala Leu Ala Lys Gln Asn Lys Leu Gln Gly Asn Ala Ala Val Pro Leu  
85 90 95  
Ile Pro Lys Thr Ala Asn Ala  
100

<210> 252

<211> 94

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..94:Ceres Seq. ID 1448187

<400>252

Met Ile Ala Thr Gln Glu Glu Met Ser Ala Ala Lys Ile Ala Leu Gly  
1 5 10 15  
Ser Arg Asp Met Cys Ala His Leu Leu Ile Pro Leu Asn Lys Cys Arg  
20 25 30  
Gln Ala Glu Phe Tyr Leu Pro Trp Lys Cys Glu Asp Glu Arg His Val  
35 40 45  
Tyr Glu Lys Cys Glu Tyr Glu Leu Val Met Glu Arg Met Leu Ala Met  
50 55 60  
Lys Lys Ile Arg Glu Glu Ala Leu Ala Lys Gln Asn Lys Leu Gln  
65 70 75 80  
Gly Asn Ala Ala Val Pro Leu Ile Pro Lys Thr Ala Asn Ala  
85 90

<210> 253

<211> 87

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..87:Ceres Seq. ID 1448188

<400>253

Met Ser Ala Ala Lys Ile Ala Leu Gly Ser Arg Asp Met Cys Ala His  
1 5 10 15  
Leu Leu Ile Pro Leu Asn Lys Cys Arg Gln Ala Glu Phe Tyr Leu Pro  
20 25 30  
Trp Lys Cys Glu Asp Glu Arg His Val Tyr Glu Lys Cys Glu Tyr Glu  
35 40 45  
Leu Val Met Glu Arg Met Leu Ala Met Lys Lys Ile Arg Glu Glu  
50 55 60  
Ala Leu Ala Lys Gln Asn Lys Leu Gln Gly Asn Ala Ala Val Pro Leu  
65 70 75 80  
Ile Pro Lys Thr Ala Asn Ala  
85

<210> 254

<211> 2034

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..2034:Ceres Seq. ID 1450875

<400>254

tttatcaaaa tcgaggtctg ctcttagatt gtgttcgagc gggcgatag ctgtgcctt  
cacttcaatg tcaatgctc agaaccgtgc ttacgcgat tcttcgat tccggttcc  
ttttttct tcttctct ctcctctcc gtacgattct ccggcaatc aaattttac  
aaactctagt aagtcctaaag cagagcctga tgagcctaaa ggatcggtt ttgactctga  
60 120 180 240

ggctcttgag agagctgcta aagctcttag agatataat agctctcccc attccaacaa 300  
ggtgtttgat ctcatagga agcagagaaa aactcgggtta gcrgaattag cgcagagagc 360  
ttccattcac gaagctcttc aagcaacaaa tgatatttgc agacacagaa aatggcttga 420  
ggacagagga aactttttgc agacacagcg gaaataccaaa ggcgaataat tgcataatga 480  
ggatgaatg gcagaaaga gacacagagc agatcatgaa gctcagagggc atcataatgt 540  
ggaatgggt aagatgaag aggcgtcttc tatcagaaa gagaagggaa aaatcgccac 600  
agaagaacag atccaagccc agcatcgcca aactgagaaa gagaagctg aacttgccag 660  
agagacatt cgtgtcaag ccatgtcgga agtgaeggt cgggtctcatg agaccaagt 720  
tacgaagag caaacagaa gattgttat ggaagagatt aatgtgaaa gagagaagt 780  
gctgtgcta atcaacaaa tgttcagcca catcgaagg ggaatcagga cctattaac 840  
tgatcaaat aagctgatta tgactgttgg aggagctact gcattagctg cagggttta 900  
tacaactgt gaaggagcta gaattacatg gggttatat aatagatgc ttgacaacc 960  
gtaactatt cgaatactt ccatggttag atttccatgg acaggtcag tgtctcaatt 1020  
taagaacag atctcagggg ctgcagcgc ttctgcagca gaaggcaaaa agccgcttga 1080  
taagttaatt ctcataact ctttgaaga acgaatcgag cgtctcgcta gagctacagc 1140  
aaacacaaa tccatacag cacatctcg caactgatg tttttggac ctctgtgac 1200  
cgttaaaact atgttggcaa ggggaatagc toggaaatcg ggtctggatt atcetatgat 1260  
tgatggggt gatgtgtgc ccttggatc acaggtgtt actaaatcc atcgatatatt 1320  
tgatggggt aagaatcga acagaggtt actcctttc attgatgag cagatgcttt 1380  
tcataggcg gctaacagca ctacatagag tgaggtcaa cgtagtgc tgaaagcttt 1440  
gctttccga actgggtgac aatctcgga catgtcttt gcttggcta caaacagacc 1500  
tgagatctc gatagtgcg ttacagacag gattgatgaa gtcattgagt tccacttcc 1560  
agggaaagaa gaactgttca agcttcaa tctctatc aacaataatc taagatggg 1620  
tgataaac gaagacaaa aaccgaatg gatcatitg titaagaagc tctcacagaa 1680  
gataccgtt gaagagact taactgata agtgatttct gagggtcaa agagaagaa 1740  
agattctt ggcgtaga ttcaagct tgtgttgga gtacaagct gagggtacgg 1800  
acagcgat tgtgttttg attccagct ttttaagag attgtgat ataaagttga 1860  
agacacac gaaagcaka tgbtcttc tgaaagttt cagccattac tctctctta 1920  
gtcatgtga tttgtatt acatgtact tacttgatca gaagaagaa tctgtatt 1980  
ttaaaatgag tcttaaaatt gaatttttag attaaacatg ttaagaagt ttac 2034

&lt;210&gt; 255

&lt;211&gt; 639

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..639:Ceres Seq. ID 1450876

<400>255  
Leu Ser Lys Ser Met Ala Ala Ser Arg Leu Cys Ser Ala Ala Ala Ile  
1 5 10 15  
Ala Ala Ala Phe Thr Ser Met Ser Met Ser Gln Asn Arg Ala Tyr Ala  
20 25 30  
Asp Ser Ser Arg Phe Arg Phe Pro Phe Phe Ser Ser Ser Pro Ser Pro  
35 40 45  
Pro Pro Ser Asp Ser Pro Ala Asn Gln Ser Ser Thr Asn Ser Ser Lys  
50 55 60  
Ser Lys Ala Glu Pro Asp Glu Pro Lys Gly Ser Gly Phe Asp Pro Glu  
65 70 75 80  
Ala Leu Glu Arg Ala Ala Lys Ala Leu Arg Asp Ile Asn Ser Ser Pro  
85 90 95  
His Ser Lys Gln Val Phe Asp Leu Met Arg Lys Gln Glu Lys Thr Arg  
100 105 110  
Leu Ala Glu Leu Ala Ala Glu Thr Ser His Tyr Glu Ala Ile Gln Ala  
115 120 125  
His Asn Asp Ile Gly Arg Gln Gln Lys Leu Ala Glu Asp Gln Arg Asn  
130 135 140  
Leu Leu Gln Thr Gln Ala Gln Thr Lys Ala Gln Asn Leu Arg Tyr Glu

145 150 155 160  
Asp Glu Leu Ala Arg Lys Arg Gln Thr Asp His Glu Ala Gln Arg  
165 170 175  
His His Asn Val Glu Leu Val Lys Met Gln Glu Ala Ser Ser Ile Arg  
180 185 190  
Lys Glu Lys Ala Lys Ile Ala Thr Glu Gln Ile Gln Ala Gln His  
195 200 205  
Arg Gln Thr Glu Lys Glu Arg Ala Glu Leu Glu Arg Glu Thr Ile Arg  
210 215 220  
Val Lys Ala Met Ala Glu Ala Glu Gly Arg Ala His Glu Ala Lys Leu  
225 230 235  
Thr Glu Glu Gln Asn Arg Arg Leu Met Glu Arg Ile Asn Gly Glu  
240 245 250 255  
Arg Glu Lys Thr Leu Ala Ala Ile Asn Thr Met Phe Ser His Ile Glu  
260 265 270  
Gly Gly Phe Arg Thr Leu Thr Asp Arg Asn Lys Leu Ile Met Thr  
275 280 285  
Val Gly Gly Ala Thr Ala Leu Ala Ala Gly Val Tyr Thr Arg Glu  
290 295 300  
Gly Ala Arg Val Thr Trp Gly Tyr Ile Asn Arg Met Leu Gly Gln Pro  
305 310 315 320  
Ser Leu Ile Arg Glu Ser Ser Met Arg Arg Phe Pro Trp Thr Gly Ser  
325 330 335  
Val Ser Gln Phe Lys Asn Arg Ile Ser Gly Ala Ala Ala Ser Ala  
340 345 350  
Ala Glu Gly Lys Lys Pro Leu Asp Asn Val Ile Leu His Thr Ser Leu  
355 360 365  
Lys Lys Arg Ile Glu Arg Leu Ala Arg Ala Thr Ala Asn Thr Lys Ser  
370 375 380  
His Gln Ala Pro Phe Arg Asn Met Met Phe Tyr Gly Pro Pro Gly Thr  
385 390 395 400  
Gly Lys Thr Met Val Ala Arg Glu Ile Ala Arg Lys Ser Gly Leu Asp  
405 410 415  
Tyr Ala Met Met Thr Gly Gly Asp Val Ala Pro Leu Gly Ser Gln Ala  
420 425 430  
Val Thr Lys Ile His Gln Ile Phe Asp Trp Ala Lys Lys Ser Asn Arg  
435 440 445  
Gly Leu Leu Leu Phe Ile Asp Glu Ala Asp Ala Phe Leu Cys Glu Arg  
450 455 460  
Asn Ser Thr Tyr Met Ser Glu Ala Gln Arg Ser Ala Leu Asn Ala Leu  
465 470 475 480  
Leu Phe Arg Thr Gly Asp Gln Ser Arg Asp Ile Val Leu Val Leu Ala  
485 490 495  
Thr Asn Arg Pro Gly Asp Leu Asp Ser Ala Val Thr Asp Arg Ile Asp  
500 505 510  
Glu Val Ile Glu Phe Pro Leu Pro Gly Glu Glu Arg Phe Lys Leu  
515 520 525  
Leu Asn Leu Tyr Leu Asn Lys Tyr Leu Lys Met Gly Asp Asn Asn Glu  
530 535 540  
Asp Thr Lys Pro Lys Trp Ser His Leu Phe Lys Lys Leu Ser Gln Lys  
545 550 555 560  
Ile Thr Val Glu Glu Asp Leu Thr Asp Lys Val Ile Ser Glu Ala Ala  
565 570 575  
Lys Lys Thr Glu Glu Phe Ser Gly Arg Glu Ile Ala Lys Leu Val Ala  
580 585 590  
Gly Val Gln Ala Gly Val Tyr Gly Arg Ala Asp Cys Val Leu Asp Ser  
595 600 605  
Gln Leu Phe Lys Glu Ile Val Glu Tyr Lys Val Glu Glu His Arg  
610 615 620  
Arg His Met Leu Ala Ser Glu Gly Phe Gln Pro Leu Leu Phe Ser  
625 630 635

<210> 256  
 <211> 635  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..635:Ceres Seq. ID 1450877  
 <400>256  
 Met Ala Ala Ser Arg Leu Cys Ser Ala Ala Ala Ile Ala Ala Phe  
 1 5 10 15  
 Thr Ser Met Ser Met Ser Gln Asn Arg Ala Tyr Ala Asp Ser Ser Arg  
 20 25 30  
 Phe Arg Phe Pro Phe Phe Ser Ser Pro Ser Pro Pro Ser Asp  
 35 40 45  
 Ser Pro Ala Asn Gln Ser Ser Thr Asn Ser Ser Lys Ser Lys Ala Glu  
 50 55 60  
 Pro Asp Glu Pro Lys Gly Ser Gly Phe Asp Pro Glu Ala Leu Glu Arg  
 65 70 75 80  
 Ala Ala Lys Ala Leu Arg Asp Ile Asn Ser Ser Pro His Ser Lys Gln  
 85 90 95  
 Val Phe Asp Leu Met Arg Lys Gln Glu Lys Thr Arg Leu Ala Glu Leu  
 100 105 110  
 Ala Ala Glu Thr Ser His Tyr Glu Ala Ile Gln Ala His Asn Asp Ile  
 115 120 125  
 Gly Arg Gln Gln Lys Leu Ala Glu Asp Gln Arg Asn Leu Leu Gln Thr  
 130 135 140  
 Gln Ala Gln Thr Lys Ala Gln Asn Leu Arg Tyr Glu Asp Glu Leu Ala  
 145 150 155 160  
 Arg Lys Arg Gln Gln Thr Asp His Glu Ala Gln Arg His Asn Val  
 165 170 175  
 Glu Leu Val Lys Met Gln Glu Ala Ser Ser Ile Arg Lys Glu Lys Ala  
 180 185 190  
 Lys Ile Ala Thr Glu Gln Ile Gln Ala Gln His Arg Gln Thr Glu  
 195 200 205  
 Lys Glu Arg Ala Glu Leu Glu Arg Glu Thr Ile Arg Val Lys Ala Met  
 210 215 220  
 Ala Glu Ala Glu Gly Arg Ala His Glu Ala Lys Leu Thr Glu Glu Gln  
 225 230 235 240  
 Asn Arg Arg Leu Leu Met Glu Arg Ile Asn Gly Glu Arg Glu Lys Trp  
 245 250 255  
 Leu Ala Ala Ile Asn Thr Met Phe Ser His Ile Glu Gly Gly Phe Arg  
 260 265 270  
 Thr Leu Leu Thr Asp Arg Asn Lys Leu Ile Met Thr Val Gly Gly Ala  
 275 280 285  
 Thr Ala Leu Ala Ala Gly Val Tyr Thr Thr Arg Glu Gly Ala Arg Val  
 290 295 300  
 Thr Trp Gly Tyr Ile Asn Arg Met Leu Gly Gln Pro Ser Leu Ile Arg  
 305 310 315 320  
 Glu Ser Ser Met Arg Arg Phe Pro Trp Thr Gly Ser Val Ser Gln Phe  
 325 330 335  
 Lys Asn Arg Ile Ser Gly Ala Ala Ala Ser Ala Ala Glu Gly Lys  
 340 345 350  
 Lys Pro Leu Asp Asn Val Ile Leu His Thr Ser Leu Lys Lys Arg Ile  
 355 360 365  
 Glu Arg Leu Ala Arg Ala Thr Ala Asn Thr Lys Ser His Gln Ala Pro  
 370 375 380  
 Phe Arg Asn Met Met Phe Tyr Gly Pro Pro Gly Thr Gly Lys Thr Met

385 390 395 400  
 Val Ala Arg Glu Ile Ala Arg Lys Ser Gly Leu Asp Tyr Ala Met Met  
 405 410 415  
 Thr Gly Gly Asp Val Ala Pro Leu Gly Ser Gln Ala Val Thr Lys Ile  
 420 425 430  
 His Gln Ile Phe Asp Trp Ala Lys Lys Ser Asn Arg Gly Leu Leu Leu  
 435 440 445  
 Phe Ile Asp Glu Ala Asp Ala Phe Leu Cys Glu Arg Asn Ser Thr Tyr  
 450 455 460  
 Met Ser Glu Ala Gln Arg Ser Ala Leu Asn Ala Leu Leu Phe Arg Thr  
 465 470 475  
 Gly Asp Gln Ser Arg Asp Ile Val Leu Val Leu Ala Thr Asn Arg Pro  
 485 490 495  
 Gly Asp Leu Asp Ser Ala Val Thr Asp Arg Ile Asp Glu Val Ile Glu  
 500 505 510  
 Phe Pro Leu Pro Gly Glu Glu Arg Phe Lys Leu Leu Asn Leu Tyr  
 515 520 525  
 Leu Asn Lys Tyr Leu Lys Met Gly Asp Asn Asn Glu Asp Thr Lys Pro  
 530 535 540  
 Lys Trp Ser His Leu Phe Lys Lys Leu Ser Gln Lys Ile Thr Val Glu  
 545 550 555 560  
 Glu Asp Leu Thr Asp Lys Val Ile Ser Glu Ala Ala Lys Lys Thr Glu  
 565 570 575  
 Gly Phe Ser Gly Arg Glu Ile Ala Lys Leu Val Ala Gly Val Gln Ala  
 580 585 590  
 Gly Val Tyr Gly Arg Ala Asp Cys Val Leu Asp Ser Gln Leu Phe Lys  
 595 600 605  
 Glu Ile Val Glu Tyr Lys Val Glu Glu His His Arg Arg His Met Leu  
 610 615 620  
 Ala Ser Glu Gly Phe Gln Pro Leu Leu Phe Ser  
 625 630 635  
 <210> 257  
 <211> 617  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..617:Ceres Seq. ID 1450878  
 <400>257  
 Met Ser Met Ser Gln Asn Arg Ala Tyr Ala Asp Ser Ser Arg Phe Arg  
 1 5 10 15  
 Phe Pro Phe Phe Ser Ser Pro Ser Pro Pro Pro Ser Asp Ser Pro  
 20 25 30  
 Ala Asn Gln Ser Ser Thr Asn Ser Ser Lys Ser Lys Ala Glu Pro Asp  
 35 40 45  
 Glu Pro Lys Gly Ser Gly Phe Asp Pro Glu Ala Leu Glu Arg Ala Ala  
 50 55 60  
 Lys Ala Leu Arg Asp Ile Asn Ser Ser Pro His Ser Lys Gln Val Phe  
 65 70 75 80  
 Asp Leu Met Arg Lys Gln Glu Lys Thr Arg Leu Ala Glu Leu Ala Ala  
 85 90 95  
 Glu Thr Ser His Tyr Glu Ala Ile Gln Ala His Asn Asp Ile Gly Arg  
 100 105 110  
 Gln Gln Lys Leu Ala Glu Asp Gln Arg Asn Leu Leu Gln Thr Gln Ala  
 115 120 125  
 Gln Thr Lys Ala Gln Asn Leu Arg Tyr Glu Asp Glu Leu Ala Arg Lys  
 130 135 140

Arg Gln Thr Asp His Glu Ala Gln Arg His Asn Val Glu Leu  
 145 150 155  
 Val Lys Met Gln Glu Ala Ser Ile Arg Lys Glu Lys Ala Lys Ile  
 165 170 175  
 Ala Thr Glu Glu Ile Gln Ala Gln His Arg Gln Thr Glu Lys Glu  
 180 185  
 Arg Ala Glu Leu Glu Arg Glu Thr Ile Arg Val Lys Ala Met Ala Glu  
 195 200 205  
 Ala Glu Gly Arg Ala His Glu Ala Lys Leu Thr Glu Glu Gln Asn Arg  
 210 215 220  
 Arg Leu Leu Met Glu Arg Ile Asn Gly Glu Arg Glu Lys Trp Leu Ala  
 225 230 235 240  
 Ala Ile Asn Thr Met Phe Ser His Ile Glu Gly Gly Phe Arg Thr Leu  
 245 250 255  
 Leu Thr Asp Arg Asn Lys Leu Ile Met Thr Val Gly Gly Ala Thr Ala  
 260 265 270  
 Leu Ala Ala Gly Val Tyr Thr Thr Arg Glu Gly Ala Arg Val Thr Trp  
 275 280 285  
 Gly Tyr Ile Asn Arg Met Leu Gly Gln Pro Ser Leu Ile Arg Glu Ser  
 290 295 300  
 Ser Met Arg Arg Phe Pro Trp Thr Gly Ser Val Ser Gln Phe Lys Asn  
 305 310 315 320  
 Arg Ile Ser Gly Ala Ala Ala Ser Ala Ala Glu Gly Lys Lys Pro  
 325 330 335  
 Leu Asp Asn Val Ile Leu His Thr Ser Leu Lys Lys Arg Ile Glu Arg  
 340 345 350  
 Leu Ala Arg Ala Thr Ala Asn Thr Lys Ser His Gln Ala Pro Phe Arg  
 355 360 365  
 Asn Met Met Phe Tyr Gly Pro Pro Gly Thr Gly Lys Thr Met Val Ala  
 370 375 380  
 Arg Glu Ile Ala Arg Lys Ser Gly Leu Asp Tyr Ala Met Met Thr Gly  
 385 390 395 400  
 Gly Asp Val Ala Pro Leu Gly Ser Gln Ala Val Thr Lys Ile His Gln  
 405 410 415  
 Ile Phe Asp Trp Ala Lys Lys Ser Asn Arg Gly Leu Leu Phe Ile  
 420 425 430  
 Asp Glu Ala Asp Ala Phe Leu Cys Glu Arg Asn Ser Thr Tyr Met Ser  
 435 440 445  
 Glu Ala Gln Arg Ser Ala Leu Asn Ala Leu Leu Phe Arg Thr Gly Asp  
 450 455 460  
 Gln Ser Arg Asp Ile Val Leu Val Leu Ala Thr Asn Arg Pro Gly Asp  
 465 470 475 480  
 Leu Asp Ser Ala Val Thr Asp Arg Ile Asp Glu Val Ile Glu Phe Pro  
 485 490 495  
 Leu Pro Gly Glu Glu Glu Arg Phe Lys Leu Leu Asn Leu Tyr Leu Asn  
 500 505 510  
 Lys Tyr Leu Lys Met Gly Asp Asn Asn Glu Asp Thr Lys Pro Lys Trp  
 515 520 525  
 Ser His Leu Phe Lys Lys Leu Ser Gln Lys Ile Thr Val Glu Glu Asp  
 530 535 540  
 Leu Thr Asp Lys Val Ile Ser Glu Ala Ala Lys Lys Thr Glu Gly Phe  
 545 550 555 560  
 Ser Gly Arg Glu Ile Ala Lys Leu Val Ala Gly Val Gln Ala Gly Val  
 565 570 575  
 Tyr Gly Arg Ala Asp Cys Val Leu Asp Ser Gln Leu Phe Lys Glu Ile  
 580 585 590  
 Val Glu Tyr Lys Val Glu Glu His His Arg Arg His Met Leu Ala Ser  
 595 600 605  
 Glu Gly Phe Gln Pro Leu Leu Phe Ser  
 610 615

<210> 258  
 <211> 478  
 <212> DNA  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..478:Ceres Seq. ID 1459191  
 <400>258  
 aacctcttg gaagagtct caacacttc agagaaaaag aacaaggaag atcccgaaa  
 atggcaagg cgaattacg ttcagctct tcccgacag tgaactcgc agtcgaga  
 catcgctgc tctaagcga aactttctt ettcgcgag ccatgacgat gctatgaag  
 ctgcgaatg ggagaagata acttatcgg gtattgctag ttgactgct ctactgtct  
 atttttatc caaggccat catcaaggc agacaagga gtittcttgg ggtccgagt  
 gtctgttga gdtgaagcac aacaagagc actgactctt gctggtcat aatacgtct  
 tctagtitta ttgaaagc taaatgttt taccgtatt gtctcacg ttgtcacg  
 attgtctact ccaatctct ttccttggg gggaataaaa agttaataact ttgcttg  
 60 120 180 240 300 360 420 478  
 <210> 259  
 <211> 66  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..66:Ceres Seq. ID 1459192  
 <400>259  
 Asn Pro Leu Gly Lys Ser Leu Asn Thr Cys Arg Glu Lys Glu Gln Gly  
 1 5 10 15  
 Arg Ser Arg Lys Met Ala Thr Ala Ile Val Arg Ser Ala Leu Ser Arg  
 20 25 30  
 Ala Val Thr Arg Ala Val Arg Arg His Pro Ser Leu Leu Ser Glu Thr  
 35 40 45  
 Phe Pro Leu Pro Pro Ala Met Thr Met Leu Met Lys Leu Arg Ser Gly  
 50 55 60  
 Arg Arg  
 65  
 <210> 260  
 <211> 110  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..110:Ceres Seq. ID 1459193  
 <400>260  
 Pro Ser Trp Lys Glu Ser Gln His Leu Gln Arg Lys Arg Thr Arg Lys  
 1 5 10 15  
 Ile Pro Glu Asn Gly Asn Gly Asp Cys Thr Phe Ser Ser Phe Pro Ser  
 20 25 30  
 Ser Asp Ser Arg Ser Pro Lys Thr Ser Val Ala Pro Lys Arg Asn Phe  
 35 40 45  
 Ser Ser Ser Ala Gly His Asp Asp Ala Tyr Glu Ala Ala Lys Trp Glu  
 50 55 60



Lys Ile Thr Tyr Leu Gly Ile Ala Ser Cys Thr Ala Leu Ala Val Tyr  
65 70 75 80  
Val Leu Ser Lys Gly His His Gly Glu Asp Lys Glu Phe Pro Trp  
85 90 95  
Gly Pro Asp Gly Leu Phe Glu Val Lys His Asn Lys Glu His  
100 105 110

&lt;210&gt; 261

&lt;211&gt; 741

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..741:Ceres Seq. ID 1461848

&lt;400&gt;261

attttagtac attgttgacc attcttttcg tatagactac tatctctgat ctcttcgag  
120 ttaagtcagt aactagaaa attcagaagc gctctcaatc tcaaaaatat ccatggcggc  
180 gattacagaa ttctaccgaa aagagtaagg atattctgtt ctgcctctg tcttctactg  
240 ttctctaac ctgtgatgg gtgtccaagt cggcagagct cgcgaaaggt acaactccc  
300 gtatccaact ctatgtgcaa tagaatcaga aacaagaat gctaagctct tcaactggt  
360 tcagagaga catcaaaact ctttagagat gatgccaatg tatttctatc tgatgatcct  
420 cggfgggatg aagcaccttt gtatctgtac tggccttggt ttgctttaca acgttagccg  
480 attctctac ttcaaggtt atgtactggt agatgccatg aagcgtctta cgtatcgga  
540 atacgttttc ttgggttgc taggtctgat gatatgacc atctcgtttg gtgtcactct  
600 gatccttgt taagctactc gtttctggg ttaatgattc tctgtttgc tcaagaata  
660 tagaaccat gcttgaagc tgtccacaaa acttgtgtaa tcttttagag ttgtcactt  
720 ttaaaagttt ttaataaac atggcttcatt agaacagttg aaatttcaca tccgtagag  
741 ttaataaga ttggaattat g

&lt;210&gt; 262

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..146:Ceres Seq. ID 1461849

&lt;400&gt;262

Met Ala Ala Ile Thr Glu Phe Leu Pro Lys Glu Tyr Gly Tyr Val Val  
1 5 10 15  
Leu Val Leu Val Phe Tyr Cys Phe Leu Asn Leu Trp Met Gly Ala Gln  
20 25 30  
Val Gly Arg Ala Arg Lys Arg Tyr Asn Val Pro Tyr Pro Thr Tyr  
35 40 45  
Ala Ile Glu Ser Glu Asn Lys Asp Ala Lys Leu Phe Asn Cys Val Gln  
50 55 60  
Arg Gly His Gln Asn Ser Leu Glu Met Met Pro Met Tyr Phe Ile Leu  
65 70 75 80  
Met Ile Leu Gly Gly Met Lys His Pro Cys Ile Cys Thr Gly Leu Gly  
85 90  
Leu Leu Tyr Asn Val Ser Arg Phe Tyr Phe Lys Gly Tyr Ala Thr  
100 105 110  
Gly Asp Pro Met Lys Arg Leu Thr Ile Gly Lys Tyr Gly Phe Leu Gly  
115 120  
Leu Leu Gly Leu Met Ile Cys Thr Ile Ser Phe Gly Val Thr Leu Ile  
130 140

125

SUBSTITUTE SHEET (RULE 26)

Leu Ala  
145

&lt;210&gt; 263

&lt;211&gt; 118

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..118:Ceres Seq. ID 1461850

&lt;400&gt;263

Met Gly Ala Gln Val Gly Arg Ala Arg Lys Arg Tyr Asn Val Pro Tyr  
1 5 10 15  
Pro Thr Leu Tyr Ala Ile Glu Ser Glu Asn Lys Asp Ala Lys Leu Phe  
20 25 30  
Asn Cys Val Gln Arg Gly His Gln Asn Ser Leu Glu Met Met Pro Met  
35 40 45  
Tyr Phe Ile Leu Met Ile Leu Gly Gly Met Lys His Pro Cys Ile Cys  
50 55 60  
Thr Gly Leu Gly Leu Leu Tyr Asn Val Ser Arg Phe Phe Tyr Phe Lys  
65 70 75 80  
Gly Tyr Ala Thr Gly Asp Pro Met Lys Arg Leu Thr Ile Gly Lys Tyr  
85 90 95  
Gly Phe Leu Gly Leu Gly Leu Met Ile Cys Thr Ile Ser Phe Gly  
100 105 110  
Val Thr Leu Ile Leu Ala  
115

&lt;210&gt; 264

&lt;211&gt; 74

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..74:Ceres Seq. ID 1461851

&lt;400&gt;264

Met Met Pro Met Tyr Phe Ile Leu Met Ile Leu Gly Gly Met Lys His  
1 5 10 15  
Pro Cys Ile Cys Thr Gly Leu Gly Leu Leu Tyr Asn Val Ser Arg Phe  
20 25 30  
Phe Tyr Phe Lys Gly Tyr Ala Thr Gly Asp Pro Met Lys Arg Leu Thr  
35 40 45  
Ile Gly Lys Tyr Gly Phe Leu Gly Leu Leu Gly Leu Met Ile Cys Thr  
50 55 60  
Ile Ser Phe Gly Val Thr Leu Ile Leu Ala  
65 70

&lt;210&gt; 265

&lt;211&gt; 469

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

126

SUBSTITUTE SHEET (RULE 26)

<223> LOCATION: 1..469:Ceres Seq. ID 1472772

<400>265  
 atatagaata taaccatatt ccgacacatga agatcgtgac attgtactc gtgcgtctcg 60  
 tcaactattc gacatcattc ccgctgcga tcaagccga agacacggga gatacaggaa 120  
 atgtggaggt gacatgtgac gcaaggcagc ttcagcttgg ccttgccggg attacggggg 180  
 gaggacaacc ctgggtgca tgggtgcga agcttacaga tcaacagtcg tgcctatggt 240  
 gtttcgttaa gaacctgtgg ttccgacgt acattagctc tccaacgct cgcacaagtgc 300  
 tccctgtgtg caatgttgt tatcccaatt gttgaacct atctagatt tataataaaa 360  
 taaacgaag gaataaatt acattatc aaagttatg atacaattca accgtttgtg 420  
 ttaattgact ggcttacatg gttaaataa gtttaattc ttggttggc 469

<210> 266

<211> 110

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..110:Ceres Seq. ID 1472773

<400>266

Ile Glu Tyr Asn His Ile Pro Asp Met Lys Ile Val Thr Leu Val Leu 15  
 1 Val Val Phe Val Ile Leu Ser Thr Ser Phe Pro Ala Ala Ile Lys Ala 30  
 Glu Asp Thr Gly Asp Thr Gly Asn Val Gly Val Thr Cys Asp Ala Arg 45  
 35 Gln Leu Gln Pro Cys Leu Ala Ala Ile Thr Gly Gly Gly Gln Pro Ser 60  
 50 Gly Ala Cys Cys Ala Lys Leu Thr Glu Gln Gln Ser Cys Leu Cys Gly 75  
 65 Phe Ala Lys Asn Pro Ala Phe Ala Gln Tyr Ile Ser Ser Pro Asn Ala 90  
 85 Arg Lys Val Leu Leu Ala Cys Asn Val Ala Tyr Pro Thr Cys 105  
 100

<210> 267

<211> 102

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..102:Ceres Seq. ID 1472774

<400>267  
 Met Lys Ile Val Thr Leu Val Leu Val Phe Val Ile Leu Ser Thr 15  
 1 Ser Phe Pro Ala Ala Ile Lys Ala Glu Asp Thr Gly Asp Thr Gly Asn 30  
 20 Val Gly Val Thr Cys Asp Ala Arg Gln Leu Gln Pro Cys Leu Ala Ala 45  
 35 Ile Thr Gly Gly Gln Pro Ser Gly Ala Cys Cys Ala Lys Leu Thr 60  
 50 Gln Gln Gln Ser Cys Leu Cys Gly Phe Ala Lys Asn Pro Ala Phe Ala 75  
 65 Gln Tyr Ile Ser Ser Pro Asn Ala Arg Lys Val Leu Leu Ala Cys Asn 90  
 85

127

SUBSTITUTE SHEET (RULE 26)

Val Ala Tyr Pro Thr Cys  
 100

<210> 268

<211> 1056

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..1056:Ceres Seq. ID 1533352

<400>268

aaaaaaaa accaaagcaa aaaaatggct ttgaactca cttccgcg ttcagtttc 60  
 tcaaatcaa ggagattatc ttctcttcg ttaattccga taaggtcaaa atccacattc 120  
 accgatttc gatcgagaac cgg-gtttat ttaagcaaaa cgaaggcgct tcagtcgtct 180  
 acaaaactga gtgtggcggc ggagatcct gcggcgacaa ttgcgacgga tgattggggg 240  
 aaagtatcgg cggttctggt tgatatggac ggtgtgctt gtaacagtga agatctttct 300  
 agacgcgcgg ccgtggatgt ttttacggag atgggagtg aagtcactgt ggacgatttc 360  
 gttccttta tgggaacagg tgaagocaa ttittagag gtgtgtcttc agtcaagaa 420  
 gttaaagat ttgatccaga tgcagctaaa aagagattct ttgaatatata tctcgataag 480  
 tatggaagc cagaatctgg gattggattt ccaggagcat tggagcttgt tactgagtgt 540  
 aagaacaag gccttaaggt cgtgtgttga tctatgtcgt accgtatcaa agtgaatgg 600  
 aatctgaag cgtcgtggtt gtccttgacc atgtttgatg cttatgttcc agcagatgcc 660  
 ttgagaatt tgaaccagc tccagatatt ttctggcgtg ctgcaaatgt cttagtgttg 720  
 cctaccagc agtgtgttgt tatgaagat gcgtgtcgtg gagtccaagc cgcacaagct 780  
 gcgaactga gattatagc cgtaaaaact actttatcgt aagcaattct taaggatgct 840  
 ggctctcaa tgatcagaga cgtatatgga aacatctcaa tcaatgacat tctcaactgt 900  
 ggctcagtt ctaccagat gtatgtcaa agaaattcga tggaaatat cgtctcttc 960  
 atgtatatt tattcttgt ttactccttt tgaaccttt tgaataagg ggtcttcttc 1020  
 gtaacgagat tacacattta aacaaatctt tctctgt 1056

<210> 269

<211> 307

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..307:Ceres Seq. ID 1533353

<400>269

Lys Lys Lys Asn Gln Ser Lys Lys Met Ala Leu Lys Leu Thr Ser Pro 15  
 1 Pro Ser Val Phe Ser Gln Ser Arg Arg Leu Ser Ser Ser Leu Ile 30  
 20 Pro Ile Arg Ser Lys Ser Thr Phe Thr Gly Phe Arg Ser Arg Thr Gly 45  
 35 Val Tyr Leu Ser Lys Thr Thr Ala Leu Gln Ser Ser Thr Lys Leu Ser 60  
 50 Val Ala Ala Glu Ser Pro Ala Ala Thr Ile Ala Thr Asp Asp Trp Gly 80  
 65 Lys Val Ser Ala Val Leu Phe Asp Met Asp Gly Val Leu Cys Asn Ser 95  
 85 Glu Asp Leu Ser Arg Ala Ala Val Asp Val Phe Thr Glu Met Gly 110  
 100 Val Glu Val Thr Val Asp Asp Phe Val Pro Phe Met Gly Thr Gly Glu 125  
 115 Ala Lys Phe Leu Gly Gly Val Ala Ser Val Lys Glu Val Lys Gly Phe 156

128

SUBSTITUTE SHEET (RULE 26)

130  
 Asp Pro Asp Ala Ala Lys Lys Arg Phe Phe Glu Ile Tyr Leu Asp Lys  
 145 150 155  
 Tyr Ala Lys Pro Glu Ser Gly Ile Gly Phe Pro Gly Ala Leu Glu Leu  
 165 170 175  
 Val Thr Glu Cys Lys Asn Lys Gly Ileu Lys Val Ala Val Ala Ser Ser  
 180 185 190  
 Ala Asp Arg Ile Lys Val Asp Ala Asn Leu Lys Ala Ala Gly Leu Ser  
 195 200 205  
 Leu Thr Met Phe Asp Ala Ile Val Ser Ala Asp Ala Phe Glu Asn Leu  
 210 215 220  
 Lys Pro Ala Pro Asp Ile Phe Leu Ala Ala Ala Lys Ile Leu Gly Val  
 225 230 235  
 Pro Thr Ser Glu Cys Val Val Ile Glu Asp Ala Leu Ala Gly Val Gln  
 245 250 255  
 Ala Ala Gln Ala Ala Asn Met Arg Cys Ile Ala Val Lys Thr Thr Leu  
 260 265 270  
 Ser Glu Ala Ile Leu Lys Asp Ala Gly Pro Ser Met Ile Arg Asp Asp  
 275 280 285  
 Ile Gly Asn Ile Ser Ile Asn Asp Ile Leu Thr Gly Gly Ser Asp Ser  
 290 295 300  
 Thr Ser Met  
 305

<210> 270  
 <211> 299  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..299:Ceres Seq. ID 1533354

<400>270  
 Met Ala Leu Lys Leu Thr Ser Pro Pro Ser Val Phe Ser Gln Ser Arg  
 1 5 10 15  
 Arg Leu Ser Ser Ser Leu Ile Pro Ile Arg Ser Lys Ser Thr Phe  
 20 25 30  
 Thr Gly Phe Arg Ser Arg Thr Gly Val Tyr Leu Ser Lys Thr Thr Ala  
 35 40 45  
 Leu Gln Ser Ser Thr Lys Leu Ser Val Ala Ala Glu Ser Pro Ala Ala  
 50 55 60  
 Thr Ile Ala Thr Asp Asp Trp Gly Lys Val Ser Ala Val Leu Phe Asp  
 65 70 75  
 Met Asp Gly Val Leu Cys Asn Ser Glu Asp Leu Ser Arg Arg Ala Ala  
 85 90 95  
 Val Asp Val Phe Thr Glu Met Gly Val Glu Val Thr Val Asp Asp Phe  
 100 105 110  
 Val Pro Phe Met Gly Thr Gly Glu Ala Lys Phe Leu Gly Gly Val Ala  
 115 120 125  
 Ser Val Lys Glu Val Lys Gly Phe Asp Pro Asp Ala Ala Lys Lys Arg  
 130 135 140  
 Phe Phe Glu Ile Tyr Leu Asp Lys Tyr Ala Lys Pro Glu Ser Gly Ile  
 145 150 155  
 Gly Phe Pro Gly Ala Leu Glu Leu Val Thr Glu Cys Lys Asn Lys Gly  
 165 170 175  
 Leu Lys Val Ala Val Ala Ser Ala Asp Arg Ile Lys Val Asp Ala  
 180 185 190  
 Asn Leu Lys Ala Ala Gly Leu Ser Leu Thr Met Phe Asp Ala Ile Val  
 195 200 205

129

SUBSTITUTE SHEET (RULE 26)

Sor Ala Asp Ala Phe Glu Asn Leu Lys Pro Ala Pro Asp Ile Phe Leu  
 210 215 220  
 Ala Ala Ala Lys Ile Leu Gly Val Pro Thr Ser Glu Cys Val Val Ile  
 225 230 235  
 Glu Asp Ala Leu Ala Gly Val Gln Ala Ala Gln Ala Ala Asn Met Arg  
 245 250 255  
 Cys Ile Ala Val Lys Thr Thr Leu Ser Glu Ala Ile Leu Lys Asp Ala  
 260 265 270  
 Gly Pro Ser Met Ile Arg Asp Asp Ile Gly Asn Ile Ser Ile Asn Asp  
 275 280 285  
 Ile Leu Thr Gly Gly Ser Asp Ser Thr Ser Met  
 290 295

<210> 271  
 <211> 219  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..219:Ceres Seq. ID 1533355

<400>271  
 Met Asp Gly Val Leu Cys Asn Ser Glu Asp Leu Ser Arg Arg Ala Ala  
 1 5 10 15  
 Val Asp Val Phe Thr Glu Met Gly Val Glu Val Thr Val Asp Asp Phe  
 20 25 30  
 Val Pro Phe Met Gly Thr Gly Glu Ala Lys Phe Leu Gly Gly Val Ala  
 35 40 45  
 Ser Val Lys Glu Val Lys Gly Phe Asp Pro Asp Ala Ala Lys Lys Arg  
 50 55 60  
 Phe Phe Glu Ile Tyr Leu Asp Lys Tyr Ala Lys Pro Glu Ser Gly Ile  
 65 70 75  
 Gly Phe Pro Gly Ala Leu Glu Leu Val Thr Glu Cys Lys Asn Lys Gly  
 85 90 95  
 Leu Lys Val Ala Val Ala Ser Ser Ala Asp Arg Ile Lys Val Asp Ala  
 100 105 110  
 Asn Leu Lys Ala Ala Gly Leu Ser Leu Thr Met Phe Asp Ala Ile Val  
 115 120 125  
 Ser Ala Asp Ala Phe Glu Asn Leu Lys Pro Ala Pro Asp Ile Phe Leu  
 130 135 140  
 Ala Ala Ala Lys Ile Leu Gly Val Pro Thr Ser Glu Cys Val Val Ile  
 145 150 155  
 Glu Asp Ala Leu Ala Gly Val Gln Ala Ala Gln Ala Ala Asn Met Arg  
 165 170 175  
 Cys Ile Ala Val Lys Thr Thr Leu Ser Glu Ala Ile Leu Lys Asp Ala  
 180 185 190  
 Gly Pro Ser Met Ile Arg Asp Asp Ile Gly Asn Ile Ser Ile Asn Asp  
 195 200 205  
 Ile Leu Thr Gly Gly Ser Asp Ser Thr Ser Met  
 210 215

<210> 272  
 <211> 1420  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

130

SUBSTITUTE SHEET (RULE 26)

&lt;223&gt; LOCATION: 1..1420:Ceres Seq. ID 1534544

<400>272  
aaagattctt ggtctccgat ccagtcacat gaagattctc ggagcttctt ttgaccacac 60  
actttcaatg cggagatgctg tgaocgctca aactccatcg ctctccgagc aatatacttt 120  
ggagaagaag gtgaagaag acaagagtcg aaagcctgtt gaagtgaag aggtggcacc 180  
agaagttact acagaagctg aagaagttaa gacggagcaa gctaaggaag aactccctgt 240  
tgaagaagcg gttctgtgag ttgaagaa gttcgaactt gctctgaat caacggaggt 300  
ggctctgag gctccctgctg cagcgaga caatgctgaa gagactctg ctgctgctga 360  
agaaataat gacgaagacg ctggtgaga agttgctgaa gaaacccctg atgagatcaa 420  
gcttgadaca gctccgctgt gactcactca ggggttctt tttttttt tttttgaa 480  
ttttcacca agtctgctt cagcgactg ctgtccctcc agtatgata tctccgltgc 540  
tgcagttggt tcaggcagaa attaagaag agcgtagag ttgggaaaa ctcatgtggt 600  
taggcctaaa ggggaacatc aagcaactat tgcctggta catgggttg gggacaatgg 660  
ctgagctgag tccagcttt ttgagacctt tcccttcca aatatcaat ggaattgcc 720  
gactgctct tctaaccaa taagtatt ttgtgtttt cctccacag ctgtgtttga 780  
ttgttgagc atcaatgaag atggactga tgaatgaa ggaatgagtg tggctgctgc 840  
acatgttga aatctgtgt cgaatgagcc tgcatacatt aaattagtg ttgagagatt 900  
cagcatgggt cggcgacat ctctattc tgaactgt ttgtctcg gtaaatagg 960  
aaatggcaat ccaataccta tcaatttaag cgaatcata gacttaagcg gctgctcc 1020  
ttgtgcaag acattgctg tcaactga agaggaacag atcaagaac gactgcatc 1080  
gttaccatt gtgtctgtc atggaagc tgaatgagtg gtaccgttca agttgggga 1140  
gaaatctica caggcttgc ttcaaatg gtttaagag gtgaccttca aactttacag 1200  
tgcacttggt caccacaaa tccacaga gttgagtag ttgtgcat gttgacatc 1260  
cagcctcagc ctgaagttt gatactct atgagtgc ttctgtga aaaccttca 1320  
accttgaga gtttgatga atggatgt tcaagattc acaatgtttt catgggaa 1380  
tttgtgaag acaattcgt ctagtgaac atctctct 1420

&lt;210&gt; 273

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..146:Ceres Seq. ID 1534545

<400>273  
Lys Ile Leu Gly Leu Arg Ser Ser His Leu Lys Ile Leu Gly Ala Ser 1  
1 5 10 15  
Phe Asp His Thr Leu Ser Met Ala Asp Ala Val Asn Ala Gln Thr Pro 20  
20 25 30  
Ser Leu Ser Glu Gln Tyr His Leu Glu Lys Glu Val Lys Gln Asp Thr 35  
35 40 45  
Ser Ala Lys Pro Val Glu Val Lys Glu Val Ala Pro Glu Val Thr Thr 50  
50 55 60  
Gln Ala Glu Glu Val Lys Thr Glu Gln Ala Lys Glu Glu Ser Pro Val 65  
65 70 75  
Glu Glu Ala Val Ser Val Val Glu Glu Lys Ser Glu Ser Ala Pro Glu 80  
80 85 90 95  
Ser Thr Glu Val Ala Ser Glu Ala Pro Ala Ala Ala Glu Asp Asn Ala 100  
100 105 110  
Glu Glu Thr Pro Ala Ala Glu Glu Asn Asn Asp Glu Asn Ala Ser 115  
115 120 125  
Glu Glu Val Ala Glu Glu Thr Pro Asp Glu Ile Lys Leu Glu Thr Ala 130  
130 135 140  
Pro Ala 145

&lt;210&gt; 274

&lt;211&gt; 252

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..252:Ceres Seq. ID 1534546

<400>274  
Met Ser Ile Ser Gly Ala Ala Val Gly Ser Gly Arg Asn Leu Arg Arg 1  
1 5 10 15  
Ala Val Glu Phe Gly Lys Thr His Val Val Arg Pro Lys Gly Lys His 20  
20 25 30  
Gln Ala Thr Ile Val Trp Leu His Gly Leu Gly Asp Asn Gly Ser Ser 35  
35 40 45  
Trp Ser Gln Leu Leu Glu Thr Leu Pro Leu Pro Asn Ile Lys Trp Ile 50  
50 55 60  
Cys Pro Thr Ala Pro Ser Gln Pro Ile Ser Leu Phe Gly Gly Phe Pro 65  
65 70 75  
Ser Thr Ala Trp Phe Asp Val Val Asp Ile Asn Glu Asp Gly Pro Asp 80  
80 85 90 95  
Asp Met Glu Gly Leu Asp Val Ala Ala His Val Ala Asn Leu Leu 100  
100 105 110  
Ser Asn Glu Pro Ala Asp Ile Lys Leu Gly Val Gly Gly Phe Ser Met 115  
115 120 125  
Gly Ala Ala Thr Ser Leu Tyr Ser Ala Thr Cys Phe Ala Leu Gly Lys 130  
130 135 140  
Tyr Gly Asn Gly Asn Pro Tyr Pro Ile Asn Leu Ser Ala Ile Ile Gly 145  
145 150 155 160  
Leu Ser Gly Trp Leu Pro Cys Ala Lys Thr Leu Ala Gly Lys Leu Glu 165  
165 170 175  
Glu Gln Gln Ile Lys Asn Arg Ala Ala Ser Leu Pro Ile Val Val Cys 180  
180 185 190  
His Gly Lys Ala Asp Asp Val Val Pro Phe Lys Phe Gly Glu Lys Ser 195  
195 200 205  
Ser Gln Ala Leu Leu Ser Asn Gly Phe Lys Lys Val Thr Phe Lys Pro 210  
210 215 220  
Tyr Ser Ala Leu Gly His His Thr Ile Pro Gln Glu Leu Asp Glu Leu 225  
225 230 235 240  
Cys Ala Trp Leu Thr Ser Thr Leu Ser Leu Glu Gly 250

&lt;210&gt; 275

&lt;211&gt; 155

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..155:Ceres Seq. ID 1534547

<400>275  
Met Glu Gly Leu Asp Val Ala Ala His Val Ala Asn Leu Leu Ser 1  
1 5 10 15  
Asn Glu Pro Ala Asp Ile Lys Leu Gly Val Gly Gly Phe Ser Met Gly 20  
20 25 30  
Ala Ala Thr Ser Leu Tyr Ser Ala Thr Cys Phe Ala Leu Gly Lys Tyr 35  
35 40 45  
Gly Asn Gly Asn Pro Tyr Pro Ile Asn Leu Ser Ala Ile Ile Gly Leu

50 55 60  
Ser Gly Trp Leu Pro Cys Ala Lys Thr Leu Ala Gly Lys Leu Glu Glu  
65 70 75 80  
Glu Gln Ile Lys Asn Arg Ala Ala Ser Leu Pro Ile Val Val Cys His  
85 90  
Gly Lys Ala Asp Asp Val Val Pro Phe Lys Phe Gly Glu Lys Ser Ser  
100 105 110  
Gln Ala Leu Leu Ser Asn Gly Phe Lys Lys Val Thr Phe Lys Pro Tyr  
115 120 125  
Ser Ala Leu Gly His His Thr Ile Pro Gln Glu Leu Asp Glu Leu Cys  
130 135 140  
Ala Trp Leu Thr Ser Thr Leu Ser Leu Glu Gly  
145 150 155

<210> 276  
<211> 1592  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<221> any n or Xaa = unknown

<223> LOCATION: 1..1592:Ceres Seq. ID 1567172

<400>276  
gtatgtctc tttttataac cacttctcga aaactgaac ctttgttag agaacccata 60  
gtcgtataa acattcttt tgaactgag acttggcaac ttgttttac tcaagtaa 120  
atttctgtgt ataatgtcaa agttgtgaa agttgtgcta gtctgttgg 180  
aggaagacg gcgagggcat gtgtagctg ctgtaagag cgggcaagtt gttatggc 240  
agctatgat gcctttcttt gccatgctt taagcttgc gtccactcg caaacctct 300  
tgcctaga cagcagagag tctgcttga acgtctgac gccgaaat atcgccatgc 360  
ctgcgcct caccagcca cgtggcataa ggtattaca cgtaaagtc ggaccacg 420  
tgagggcaag aagaccaca cgaatgttt tcatgatct gtccggaga tgaacgga 480  
ggataacgc gagagtcag agtgaaga gcagctcata tttaggtgc cgttgtaa 540  
ctcatggtt gagagcaat gtttaacca atccctggag aaacagaat agttccaa 600  
gagcccta agttcaaga gtagtgaga aagagatgac gacaacgtc agattgtct 660  
gaatgttg ttcccaacg acatgaact agctcagtc atgactgag ttgagactct 720  
actcgttga gggagtcag agtttcat catagaaga ctagggttg gtgagtgt 780  
aaagatgaa aagagggagg tggagaaga tgaaggagt gtgacaagag aadtgatga 840  
tcaagatgaa gttgatgaga catcccaat tgaataagc tttagctac agtacacga 900  
caagaccaca ttcatgtag gagagaaga tgaanaaga gactgatga agaatgat 960  
ggagatgga gtaaatgaga tgaatgttg gattaaaga gagagaagc agaggtct 1020  
tatgttaga ttgactatg aatcagttat ttccacttg ggagccaag ggtcccatg 1080  
gacgcacg gtccatctg aaatgaact cgaatggtt gtttccaa cccataccat 1140  
gggtgaagt ggagcagag ctcacatca caaccattc cggcgcctg ggttacact 1200  
aggagatgt gggatgag gagagagc taggtttca agataccag agaaaagag 1260  
gacaagtgt ttcccaaga agataagta caggtacgt aaattgagc cagataaag 1320  
gctccgacg aaggaaggt tgcacagag acttcaatt ggtttgctc actaagaac 1380  
ttaatcaat atgatatata attctcttg cctcaactt gctttttgt tgcatagt 1440  
ttgtgattg ttatgttct ttctgcat tcatagaga ttgtgcagt ttgtgagc 1500  
tactatga cataatata tcacaaaaa agtgaatat ctttgaaga ctgattata 1560  
tagctgatar aactgaatt ttgattgctg gt 1592

<210> 277  
<211> 407  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<221> any n or Xaa = unknown

<223> LOCATION: 1..407:Ceres Seq. ID 1567173

<400>277  
Met Met Lys Ser Leu Ala Ser Ala Val Gly Gly Lys Thr Ala Arg Ala 15  
1 Cys Asp Ser Cys Val Lys Arg Arg Ala Arg Trp Tyr Cys Ala Ala Asp 30  
20 25  
Asp Ala Phe Leu Cys His Ala Cys Asp Gly Ser Val His Ser Ala Asn 45  
35 40  
Pro Leu Ala Arg Arg His Glu Arg Val Arg Leu Lys Ser Ala Ser Ala 60  
50 55  
Gly Lys Tyr Arg His Ala Ser Pro Pro His Gln Ala Thr Trp His Gln 75  
65 70 80  
Gly Phe Thr Arg Lys Ala Arg Thr Pro Arg Gly Gly Lys Lys Ser His 95  
85 90  
Thr Met Val Phe His Asp Leu Val Pro Glu Met Ser Thr Glu Asp Gln 110  
100 105  
Ala Glu Ser Tyr Glu Val Glu Gln Leu Ile Phe Glu Val Pro Val 125  
115 120  
Met Asn Ser Met Val Glu Glu Gln Cys Phe Asn Gln Ser Leu Glu Lys 140  
130 135  
Gln Asn Glu Phe Pro Met Met Pro Leu Ser Phe Lys Ser Ser Asp Glu 150  
145 150  
Glu Asp Asp Asn Ala Glu Ser Cys Leu Asn Gly Leu Phe Pro Thr 175  
165 170  
Asp Met Glu Leu Ala Gln Phe Thr Ala Asp Val Glu Thr Leu Leu Gly 190  
Gly Gly Asp Arg Glu Phe His Ser Ile Glu Glu Leu Gly Leu Gly Glu 205  
195 200  
Met Leu Lys Ile Glu Lys Glu Glu Val Glu Glu Glu Gly Val Val 220  
210 215  
Thr Arg Glu Val His Asp Gln Asp Glu Gly Asp Glu Thr Ser Pro Phe 235  
225 230  
Glu Ile Ser Phe Asp Tyr Glu Tyr Thr His Lys Thr Thr Phe Asp Glu 250  
245 255  
Gly Glu Glu Asp Glu Lys Glu Asp Val Met Lys Asn Val Met Glu Met 270  
260 265  
Gly Val Asn Glu Met Ser Gly Gly Ile Lys Glu Glu Lys Lys Glu Lys 285  
275 280  
Ala Leu Met Leu Arg Leu Asp Tyr Glu Ser Val Ile Ser Thr Trp Gly 300  
290 295  
Gly Gln Gly Ile Pro Trp Thr Ala Arg Val Pro Ser Glu Ile Asp Leu 315  
305 310  
Asp Met Val Cys Phe Pro Thr His Thr Met Gly Glu Ser Gly Ala Glu 330  
320 335  
Ala His His Asn His Phe Arg Gly Leu Gly Leu His Leu Gly Asp 350  
340 345  
Ala Gly Asp Gly Gly Arg Glu Ala Arg Val Ser Arg Tyr Arg Glu Lys 365  
355 360  
Arg Arg Thr Arg Leu Phe Ser Lys Lys Ile Arg Tyr Glu Val Arg Lys 380  
370 375  
Leu Asn Ala Asp Lys Arg Pro Arg Met Lys Gly Arg Phe Val Lys Arg 395  
385 390 400  
Ser Ser Ile Gly Val Ala His 405

<210> 278  
<211> 406  
<212> PRT  
<213> Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..406:Ceres Seq. ID 1567174

&lt;400&gt;278

Met Lys Ser Leu Ala Ser Ala Val Gly Gly Lys Thr Ala Arg Ala Cys  
 1 5 10 15  
 Asp Ser Cys Val Lys Arg Arg Ala Arg Trp Tyr Cys Ala Ala Asp Asp  
 20 25 30  
 Ala Phe Leu Cys His Ala Cys Asp Gly Ser Val His Ser Ala Asn Pro  
 35 40 45  
 Leu Ala Arg Arg His Glu Arg Val Arg Leu Lys Ser Ala Ser Ala Gly  
 50 55 60  
 Lys Tyr Arg His Ala Ser Pro His Gln Ala Thr Trp His Gln Gly  
 65 70 75 80  
 Phe Thr Arg Lys Ala Arg Thr Pro Arg Gly Gly Lys Ser His Thr  
 85 90 95  
 Met Val Phe His Asp Leu Val Pro Glu Met Ser Thr Glu Asp Gln Ala  
 100 105 110  
 Glu Ser Tyr Glu Val Glu Glu Gln Leu Ile Phe Glu Val Pro Val Met  
 115 120 125  
 Asn Ser Met Val Glu Glu Gln Cys Phe Asn Gln Ser Leu Glu Lys Gln  
 130 135 140  
 Asn Glu Phe Pro Met Pro Leu Ser Phe Lys Ser Ser Asp Glu Glu  
 145 150 155 160  
 Asp Asp Asn Ala Glu Ser Cys Leu Asn Gly Leu Phe Pro Thr Asp  
 165 170 175  
 Met Glu Leu Ala Gln Phe Thr Ala Asp Val Glu Thr Leu Gly Gly  
 180 185 190  
 Gly Asp Arg Glu Phe His Ser Ile Glu Glu Leu Gly Leu Gly Glu Met  
 195 200 205  
 Leu Lys Ile Glu Lys Glu Glu Val Glu Glu Glu Gly Val Val Thr  
 210 215 220  
 Arg Glu Val His Asp Gln Asp Glu Gly Asp Glu Thr Ser Pro Phe Glu  
 225 230 235 240  
 Ile Ser Phe Asp Tyr Glu Tyr Thr His Lys Thr Thr Phe Asp Glu Gly  
 245 250 255  
 Glu Glu Asp Glu Lys Glu Asp Val Met Lys Asn Val Met Glu Met Gly  
 260 265 270  
 Val Asn Glu Met Ser Gly Gly Ile Lys Glu Glu Lys Lys Glu Lys Ala  
 275 280 285  
 Leu Met Leu Arg Leu Asp Tyr Glu Ser Val Ile Ser Thr Trp Gly Gly  
 290 295 300  
 Gln Gly Ile Pro Trp Thr Ala Arg Val Pro Ser Glu Ile Asp Leu Asp  
 305 310 315 320  
 Met Val Cys Phe Pro Thr His Thr Met Gly Glu Ser Gly Ala Glu Ala  
 325 330 335  
 His His His Asn His Phe Arg Gly Leu Gly Leu His Leu Gly Asp Ala  
 340 345 350  
 Gly Asp Gly Gly Arg Glu Ala Arg Val Ser Arg Tyr Arg Glu Lys Arg  
 355 360 365  
 Arg Thr Arg Leu Phe Ser Lys Lys Ile Arg Tyr Glu Val Arg Lys Leu  
 370 375 380  
 Asn Ala Asp Lys Arg Pro Arg Met Lys Gly Arg Phe Val Lys Arg Ser  
 385 390 395 400  
 Ser Ile Gly Val Ala His 405

&lt;210&gt; 279

&lt;211&gt; 310

135

SUBSTITUTE SHEET (RULE 26)

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..310:Ceres Seq. ID 1567175

&lt;400&gt;279

Met Val Phe His Asp Leu Val Pro Glu Met Ser Thr Glu Asp Gln Ala  
 1 5 10 15  
 Glu Ser Tyr Glu Val Glu Glu Gln Ile Phe Glu Val Pro Val Met  
 20 25 30  
 Asn Ser Met Val Glu Glu Gln Cys Phe Asn Gln Ser Leu Glu Lys Gln  
 35 40 45  
 Asn Glu Phe Pro Met Met Pro Leu Ser Phe Lys Ser Ser Asp Glu Glu  
 50 55 60  
 Asp Asp Asp Asn Ala Glu Ser Cys Leu Asn Gly Leu Phe Pro Thr Asp  
 65 70 75 80  
 Met Glu Leu Ala Gln Phe Thr Ala Asp Val Glu Thr Leu Leu Gly Gly  
 85 90 95  
 Gly Asp Arg Glu Phe His Ser Ile Glu Glu Leu Gly Leu Gly Glu Met  
 100 105 110  
 Leu Lys Ile Glu Lys Glu Glu Val Glu Glu Glu Gly Val Val Thr  
 115 120 125  
 Arg Glu Val His Asp Gln Asp Glu Gly Asp Glu Thr Ser Pro Phe Glu  
 130 135 140  
 Ile Ser Phe Asp Tyr Glu Tyr Thr His Lys Thr Thr Phe Asp Glu Gly  
 145 150 155 160  
 Glu Glu Asp Glu Lys Glu Asp Val Met Lys Asn Val Met Glu Met Gly  
 165 170 175  
 Val Asn Glu Met Ser Gly Gly Ile Lys Glu Glu Lys Lys Glu Lys Ala  
 180 185 190  
 Leu Met Leu Arg Leu Asp Tyr Glu Ser Val Ile Ser Thr Trp Gly Gly  
 195 200 205  
 Gln Gly Ile Pro Trp Thr Ala Arg Val Pro Ser Glu Ile Asp Leu Asp  
 210 215 220  
 Met Val Cys Phe Pro Thr His Thr Met Gly Glu Ser Gly Ala Glu Ala  
 225 230 235 240  
 His His His Asn His Phe Arg Gly Leu Gly Leu His Leu Gly Asp Ala  
 245 250 255  
 Gly Asp Gly Gly Arg Glu Ala Arg Val Ser Arg Tyr Arg Glu Lys Arg  
 260 265 270  
 Arg Thr Arg Leu Phe Ser Lys Lys Ile Arg Tyr Glu Val Arg Lys Leu  
 275 280 285  
 Asn Ala Asp Lys Arg Pro Arg Met Lys Gly Arg Phe Val Lys Arg Ser  
 290 295 300  
 Ser Ile Gly Val Ala His 305  
 310

&lt;210&gt; 280

&lt;211&gt; 520

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..520:Ceres Seq. ID 1567535

&lt;400&gt;280

136

SUBSTITUTE SHEET (RULE 26)

aaaaatccg cagaattgat taltcgtcaat tcgttaagga tctctgtggg gtttgatcga 60  
 attcatagc gaaggagaga gbaagagaga gattgaagca gtaatggcag gaacatctgg 120  
 attctcaac gcagtgaagc caaagatgat attcaggccg ccgttgatg 180  
 gggaatgcc gcgcagcgc gtccatctcg gtctgtcaa ccatttggt ggataaaga 240  
 gaacatcat gaccacccc caactgaaga gaagtggat atccacaaga tteaaacta 300  
 agtgaagat ttccagttt ctaagtgtt tacogctct ttggcgtga gctgaataa 360  
 ttgtgttt ttctgggc ttctgggtt caattctca atgttgact gatttgtct 420  
 ctcaatgtt atacaaca agaaagcac ttactcaag ttactgaaa agagaacca 480  
 tttagattt atcatgagt attatttcc atgtcaagt 520

&lt;210&gt; 281

&lt;211&gt; 57

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..57:Ceres Seq. ID 1567536

&lt;400&gt;281

Met Ala Gly Thr Ser Gly Leu Leu Asn Ala Val Lys Pro Lys Ile Gln 1  
 Thr Ile Asp Ile Gln Ala Ala Gly Trp Gly Ile Ala Ala Ala 15  
 Gly Ala Ile Trp Val Val Gln Pro Phe Gly Trp Ile Lys Lys Thr Phe 20  
 Ile Asp Pro Pro Thr Glu Glu Lys 30  
 45

&lt;210&gt; 282

&lt;211&gt; 765

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..765:Ceres Seq. ID 1569689

&lt;400&gt;282

aaccaaat ttcttgtct ctgtctcaa aatgaatca aaatctctaa agtttcaatt 60  
 ttgttgtt gttcttttt ttittaaag aatgcttca atttgcga ttgtgcttc 120  
 gccattgta ctacacaga gaaatctcaa tctcatcgc attcaaaac tccatttic 180  
 tctaactga ggtacgaag atattctcc attatctct atcgaaacc ctacagcat 240  
 cagtcagtg gtgaagcta gtggaaga ctcagattca tgcactgac tgcactgtg 300  
 tagtaagatt cagaatgtt ggataagtc tgaagatagg ttaggctta ttggttgg 360  
 ttttRgtgg taattagct ctttgggat cattaatct catcagcca atggacaaat 420  
 tgccttat ctgagcgga ttcgaacag ttggtactt gttctccag tggttacat 480  
 atcatatct ctgttctaaa ccgaacagc aggaacttc gaaattgtc aagaaatcag 540  
 tagcggaat acttgccag tgaacttgt gtgttgata atactttcat ctgtggaaga 600  
 tgatttgtt gcaagtgtt aaaaataat gacaggtgg ttgtgttc tagtcaata 660  
 atgtcatga ttgaacctt gtaactatt ttgttgtt ttgttgtt gagaacatc 720  
 aatctttt aattcaag attctcttt atgattatc gtttt 765

&lt;210&gt; 283

&lt;211&gt; 66

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

<223> any n or Xaa = unknown  
 <223> LOCATION: 1..66:Ceres Seq. ID 1569690  
 <400>283  
 Asn Gln Lys Phe Leu Arg Leu Cys Leu Gln Asn Arg Ile Lys Ile Ser 1  
 Lys Val Ser Ile Phe Leu Phe Cys Ser Phe Phe Leu Lys Asn Gly 5  
 Phe Asn Phe Cys Asn Phe Ala Phe Ala Ile Val Thr His Thr Glu Lys 10  
 Ile Gln Ser His Ile Asp Ser Lys Thr Pro Ile Phe Ser Asn Ser Arg 15  
 Tyr Glu 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55  
 60

&lt;210&gt; 284

&lt;211&gt; 107

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..107:Ceres Seq. ID 1569691

&lt;400&gt;284

Met Ala Ser Ile Ser Ala Thr Leu Pro Ser Pro Leu Leu Leu Thr Gln 1  
 Arg Lys Ser Asn Leu Thr Ser Ile Gln Lys Leu Pro Phe Ser Leu Thr 5  
 Arg Gly Thr Asn Asp Leu Ser Pro Leu Ser Leu Thr Arg Asn Pro Ser 10  
 Ser Ile Ser Leu Met Val Lys Ala Ser Gly Glu Ser Ser Asp Ser Ser 15  
 Thr Asp Leu Asp Val Val Ser Thr Thr Ile Gln Asn Val Trp Asp Lys Ser 20  
 Glu Asp Arg Leu Gly Leu Ile Gly Leu Gly Phe Xaa Trp Tyr Cys Ser 25  
 Ser Leu Gly Ile Ile Glu Ser His His Gly Asn 30  
 35  
 40  
 45  
 50  
 55  
 60  
 65  
 70  
 75  
 80  
 85  
 90  
 95  
 100

&lt;210&gt; 285

&lt;211&gt; 55

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..55:Ceres Seq. ID 1569692

&lt;400&gt;285

Met Val Lys Ala Ser Gly Glu Ser Ser Asp Ser Thr Asp Leu Asp 1  
 Val Val Ser Thr Thr Ile Gln Asn Val Trp Asp Lys Ser Glu Asp Arg Leu 5  
 Gly Leu Ile Gly Leu Gly Phe Xaa Trp Tyr Cys Ser Ser Leu Gly Ile 10  
 Ile Glu Ser His His Gly Asn 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50

&lt;210&gt; 286

&lt;211&gt; 782

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..782:Ceres Seq. ID 1571042

&lt;400&gt;286

atacaacaaa acaaaacata aaaaacaagt ggaagcttta aaacgacagg gagagagcaa 60  
 aaatggcgac gtccggaacg taactgacgg agttcccgct aaagagatcg gcagagaat 120  
 actacaagg gtgaagagac gagaaccatg tcttccctga tgcatacggc caccacatcc 180  
 aaaaattac cgttccagaa ggccgaacatg actctcacgg Gctatcagg agttggaact 240  
 acaactggga tggaaaggag gaggtgttca agagagaag agagatagac gatgagacca 300  
 aaacgttgac gttaaagaga cttgagggtc acgtgatgga gcagctcaaa gtgcacgcy 360  
 tgcctacca attacttccc aaatctgagg ataccgtgat cggcaaaatc actttaatat 420  
 gggaagagcg caacgatgat tcccagaac caagcggcta catgaattc gtcaagagct 480  
 tggctgcga catgggaac caogttagca aaacttaac atcattccca agctgctgt 540  
 catcatcac atcatcata tctcgattta taagttaaag tgttttcagt ataataaag 600  
 gggtcttgg gatcgttcac ttctatgtgt aaacogtttg gttctgtatg atgcttcgat 720  
 atattgttat gttcatgatc atatgctcgg ttogatataa tgattcttaa gattaattta 782  
 ct

&lt;210&gt; 287

&lt;211&gt; 151

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..151:Ceres Seq. ID 1571043

&lt;400&gt;287

Met Ala Thr Ser Gly Thr Tyr Val Thr Glu Val Pro Leu Lys Gly Ser 15  
 1 Ala Glu Lys Tyr Tyr Lys Arg Trp Lys Asn Glu Asn His Val Phe Pro 20  
 Asp Ala Ile Gly His His Ile Gln Asn Val Thr Val His Glu Gly Glu 35  
 35 His Asp Ser His Gly Ser Ile Arg Ser Trp Asn Tyr Thr Trp Asp Gly 40  
 50 Lys Glu Glu Val Phe Lys Glu Arg Arg Glu Ile Asp Asp Glu Thr Lys 60  
 65 Thr Leu Thr Leu Arg Gly Leu Glu Gly His Val Met Glu Gln Leu Lys 70  
 Val Tyr Asp Val Val Tyr Gln Phe Ile Pro Lys Ser Glu Asp Thr Cys 85  
 Ile Gly Lys Ile Thr Leu Ile Trp Glu Lys Arg Asn Asp Asp Ser Pro 100  
 115 Glu Pro Ser Gly Tyr Met Lys Phe Val Lys Ser Leu Val Ala Asp Met 120  
 130 Gly Asn His Val Ser Lys Thr 135  
 145

&lt;210&gt; 288

&lt;211&gt; 718

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..718:Ceres Seq. ID 1571079

&lt;400&gt;288

atticaagac atacaataa attgagtttt ttttttttaa ttgaacaaa atggggttga 60  
 gtggtttct tcatggag gttgaggtta agtctccggc tgaagagtic tgggtagccc 120  
 tgggaagg catcaatct tttcccaag ctttccctaa cgaactacaa accatccaag 180  
 tttagccgg cagcggcaac gctctcggct ccattcgctt cattacttat ggaagaagg 240  
 tctcactgt gaagtatcg ggggagagga togaagcagt gatttggag acaaaaagca 300  
 tgcatacag catcattggc ggcgaatgt tggagtacta caaaogitic aaagaaacca 360  
 tcaacttat tctaagAAC ggtgcagcc tctgaaatg gtcgttgag ttgagaaga 420  
 ccgccatga gattgatgac ccaacgtca tcaaggactt tgcgtcaag aactcaag 480  
 agatagatga gtatctcctt agcaaaacta gtgcctaca ctgaaccct taaattatat 540  
 Maagagggtt cgtctgtctc ttaagattt ttctaattaa gaagttgaat aaagtgaac 600  
 ctcttatga atatcaagt ttgtgatttc ggaatttatg cagcctagta ggcataagc 660  
 tttttacaa agccaattta gtcgaaactt ttgaaaaaat cgaacctttt ggtaaagc 718

&lt;210&gt; 289

&lt;211&gt; 155

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..155:Ceres Seq. ID 1571080

&lt;400&gt;289

Met Gly Leu Ser Gly Val Leu His Val Glu Val Glu Val Lys Ser Pro 15  
 1 Ala Glu Lys Phe Trp Val Ala Leu Gly Asp Gly Ile Asn Leu Phe Pro 20  
 Lys Ala Phe Pro Asn Asp Tyr Lys Thr Ile Gln Val Leu Ala Gly Asp 25  
 Gly Asn Ala Pro Gly Ser Ile Arg Leu Ile Thr Tyr Gly Glu Gly Ser 30  
 50 Pro Leu Val Lys Ile Ser Ala Glu Arg Ile Glu Ala Val Asp Leu Glu 60  
 65 Asn Lys Ser Met Ser Tyr Ser Ile Ile Gly Gly Glu Met Leu Glu Tyr 75  
 85 Tyr Lys Thr Phe Lys Gly Thr Ile Thr Val Ile Pro Lys Asn Gly Gly 90  
 100 Ser Leu Leu Lys Trp Ser Gly Glu Phe Glu Lys Thr Ala His Glu Ile 105  
 115 Asp Asp Pro His Val Ile Lys Asp Phe Ala Val Lys Asn Phe Lys Glu 120  
 130 Ile Asp Glu Tyr Leu Leu Lys Gln Thr Ser Ala 135  
 145

&lt;210&gt; 290

&lt;211&gt; 808

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown



<223> LOCATION: 1..808:Ceres Seq. ID 1572097

<400>290  
atacaacaa acaaacata aaaaacagt ggaagcttta aaacgagag gatagagcaa  
aaatggcag gtgggaacg tacgtgacg agttccgct aaagagatcg gccagagaa  
actacaagg gtgaagaac gagaacactg tcttccctga tgcatacgc caccacatc  
aaaatttac cgttcacga ggcgaacatg actctccag gctatcacg agtggaaat  
acacatgga tgaagaagg gaggtgttca aggaagaag agagatagac gatgagaca  
aaagctgac gttaagagg cttgaggtc acgtgatgga gcagctcaaa gttacgacg  
tcgtacaca atcattacc aaatgagg ataactgat cggcaaaatc acttaatat  
gggaagcg caacgatgt tcccagac caagcgcta catgaaatc gtcaagagt  
tggtgtga catgggaac cecgttagc CCAAAAAA AAAAAAaa ctlaatcac  
atcccaag tctgtgat catcatcac atcatcata tcatcatc catcatcac  
atcatcata tcatcatc catcatcac atcatcatc catcatcata gttaaagt  
ttcagata ataagtggg tcttggtat cgtcatc tctgtgaa cgtttggt  
ctgatgat cttcatata ttgtatgt catgatata tgcgggtc gatataatga  
ttcttaagt taattacta cacattc

<210> 291

<211> 195

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..195:Ceres Seq. ID 1572098

<400>291

Met Ala Thr Ser Gly Thr Tyr Val Thr Glu Val Pro Leu Lys Gly Ser  
1 1  
Ala Glu Lys Tyr Tyr Lys Arg. Trp Lys Asn Glu Asn His Val Phe Pro  
20 25 30  
Asp Ala Ile Gly His His Ile Gln Asn Val Thr Val His Glu Gly Glu  
35 35 40 40  
His Asp Ser His Gly Ser Ile Arg Ser Trp Asn Tyr Thr Trp Asp Gly  
50 55 60  
Lys Glu Glu Val Phe Lys Glu Arg Arg Glu Ile Asp Asp Glu Thr Lys  
65 70 75 80  
Thr Leu Thr Leu Arg Gly Leu Glu Gly His Val Met Glu Gln Leu Lys  
85 90 95  
Val Tyr Asp Val Val Tyr Gln Phe Ile Pro Lys Ser Glu Asp Thr Cys  
100 105 110  
Ile Gly Lys Ile Thr Leu Ile Trp Glu Lys Arg Asn Asp Asp Ser Pro  
115 120 125  
Glu Pro Ser Gly Tyr Met Lys Phe Val Lys Ser Leu Val Ala Asp Met  
130 135 140  
Gly Asn His Val Ser Pro Lys Lys Lys Lys Lys Leu Asn His His  
145 150 155 160  
Ser His Ser Arg Arg His His His His His His His His His His  
165 170 175  
His His His His His His His His His His His His His His Tyr  
180 185 190  
Leu Asp Leu  
195

<210> 292

<211> 104

<212> PRT

<213> Arabidopsis thaliana

141

SUBSTITUTE SHEET (RULE 26)

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..104:Ceres Seq. ID 1572099

<400>292  
Met Glu Gln Leu Lys Val Tyr Asp Val Val Tyr Gln Phe Ile Pro Lys  
1 5 10 15  
Ser Glu Asp Thr Cys Ile Gly Lys Ile Thr Leu Ile Trp Glu Lys Arg  
20 25 30  
Asn Asp Asp Ser Pro Glu Pro Ser Gly Tyr Met Lys Phe Val Lys Ser  
35 40 45  
Leu Val Ala Asp Met Gly Asn His Val Ser Pro Lys Lys Lys Lys  
50 55 60  
Lys Leu Asn His His Ser His Ser Arg Arg His His His His His  
65 70 75 80  
His His His His His His His His His His His His His His  
85 90 95  
His His His His Tyr Leu Asp Leu  
100

<210> 293

<211> 583

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..583:Ceres Seq. ID 1572890

<400>293

aaaaaaaa ataaaaaaa catgcacaa gaaataaaaa gatttgtaga atcaactaag  
aaaaatgca gcaatgat gactacattg cctcagttca atggtcttcg agccacaaa  
atcttcgag ctctgtaca agcctggca agtgttcagc caatgagacg caaggaaaat  
ggagcttgg gtcaaaagtg tgacttcac gggtcatcaa caaatctgat aatgtaacg  
tcgacgccc tgatttgtt cgcggggaga ttccgacttg cgcctacgc caataggaa  
ggacacgtg gacttagtt ggaggaacgt gactcaggtc tacaacggg tgaccggcc  
gggttcacg ttgcggcac ttgtctgtt ggacacgttg gcatatcat cgtgtagga  
gtgtccctg gcctaaaaa cattgtgct atttgagt cctaaagtc tttatttgt  
attgtaaa attgttagat tttaataca atattctat gcactgaac gagatcta  
ggatttaca agtcttatg ttattctat aatgttgat cgc

<210> 294

<211> 151

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..151:Ceres Seq. ID 1572891

<400>294

Lys Asn Lys Asn Lys Lys Asn Ile Ala Gln Glu Asn Lys Arg Phe Val  
1 5 10 15  
Glu Ser Thr Lys Lys Met Ala Ser Thr Met Met Thr Thr Leu Pro Gln  
20 25 30  
Phe Asn Gly Leu Arg Ala Thr Lys Ile Ser Ala Ala Pro Val Gln Gly  
35 40 45

142

SUBSTITUTE SHEET (RULE 26)

Leu Ala Ser Val Gln Pro Met Arg Arg Lys Gly Asn Gly Ala Leu Gly  
 50 55 60  
 Ala Lys Cys Asp Phe Ile Gly Ser Thr Asn Leu Ile Met Val Thr  
 65 70 75  
 Ser Thr Thr Leu Ile Leu Phe Ala Gly Arg Phe Gly Leu Ala Pro Ser  
 85 90 95  
 Ala Asn Arg Lys Ala Thr Ala Gly Leu Arg Leu Glu Ala Arg Asp Ser  
 100 105 110  
 Gly Leu Gln Thr Gly Asp Pro Ala Gly Phe Thr Leu Ala Asp Thr Leu  
 115 120 125  
 Ala Cys Gly Thr Val Gly His Ile Ile Gly Val Gly Val Val Leu Gly  
 130 135 140  
 Leu Lys Asn Ile Gly Ala Ile  
 145 150

&lt;210&gt; 295

&lt;211&gt; 130

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..130:Ceres Seq. ID 1572892

&lt;400&gt;295

Met Ala Ser Thr Met Met Thr Thr Leu Pro Gln Phe Asn Gly Leu Arg  
 1 5 10 15  
 Ala Thr Lys Ile Ser Ala Ala Pro Val Gln Gly Leu Ala Ser Val Gln  
 20 25 30  
 Pro Met Arg Arg Lys Gly Asn Gly Ala Ile Gly Ala Lys Cys Asp Phe  
 35 40 45  
 Ile Gly Ser Ser Thr Asn Leu Ile Met Val Thr Ser Thr Thr Leu Ile  
 50 60  
 Leu Phe Ala Gly Arg Phe Gly Leu Ala Pro Ser Ala Asn Arg Lys Ala  
 65 70 75 80  
 Thr Ala Gly Leu Arg Leu Glu Ala Arg Asp Ser Gly Leu Gln Thr Gly  
 85 90  
 Asp Pro Ala Gly Phe Thr Leu Ala Asp Thr Leu Ala Cys Gly Thr Val  
 100 105 110  
 Gly His Ile Ile Gly Val Gly Val Val Leu Gly Leu Lys Asn Ile Gly  
 115 120 125  
 Ala Ile  
 130

&lt;210&gt; 296

&lt;211&gt; 126

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..126:Ceres Seq. ID 1572893

&lt;400&gt;296

Met Met Thr Thr Leu Pro Gln Phe Asn Gly Leu Arg Ala Thr Lys Ile  
 1 5 10 15  
 Ser Ala Ala Pro Val Gln Gly Leu Ala Ser Val Gln Pro Met Arg Arg  
 20 25 30  
 Lys Gly Asn Gly Ala Leu Gly Ala Lys Cys Asp Phe Ile Gly Ser Ser

Thr Asn Leu Ile Met Val Thr Ser Thr Thr Leu Ile Leu Phe Ala Gly  
 35 40 45  
 Arg Phe Gly Leu Ala Pro Ser Ala Asn Arg Lys Ala Thr Ala Gly Leu  
 50 55 60  
 65 70 75 80  
 Arg Leu Glu Ala Arg Asp Ser Gly Leu Gln Thr Gly Asp Pro Ala Gly  
 90 95  
 Phe Thr Leu Ala Asp Thr Leu Ala Cys Gly Thr Val Gly His Ile Ile  
 100 105 110  
 Gly Val Gly Val Val Leu Gly Leu Lys Asn Ile Gly Ala Ile  
 115 120 125

&lt;210&gt; 297

&lt;211&gt; 748

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..748:Ceres Seq. ID 1573606

&lt;400&gt;297

ccttcact tcaattcat aatctctcta ctctctctc tctctaaatt caaaatggc  
 gaaactggg atgtgtgg tctttgtat cttaaggcg atagccatgg cygcaaggag  
 gactaatt ggaagaata caatgttgt tcaagtagc acttactgtg acattgcaa  
 attcgcttc gaaactcctg aatctctcta ctctatccc ggtgaacgg tgaagtacc  
 atgcaagac aggaagacaa tggagagggt ttacacagac aagctgtat cygcaaga  
 agaaagat aagttcattg tcaacagca tcaacagac cagatgtgg atgtttgt  
 tgtgaaagc tcggataaaa cctgctctaa aatctcgtt ggaagtga agtctcgt  
 gacttgaac cattacatg goattgcctc ggaatcaga catgtaaca acatgggat  
 cgagaagaa gtgagtgtg tgttctctc tctttgttt cagaagata tggttgatg  
 agatgggat gatataaaa accatctta atctctgt ttactttat gatctgtgt  
 ttcttcatt atgagtttc gagttatga agatatat ttgtattgt ttgattact  
 attgtgtc ttgagatgt ttgacttgg tgaaggata actatctgt tgttaagct  
 tcttatat tgatgtgca ttctctt

&lt;210&gt; 298

&lt;211&gt; 171

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..171:Ceres Seq. ID 1573607

&lt;400&gt;298

Met Ala Lys Leu Val Met Leu Leu Val Leu Cys Ile Leu Pro Ala Ile  
 1 5 10 15  
 Ala Met Ala Ala Arg Ser Asn Ile Gly Lys Asn Thr Met Val Val  
 20 25 30  
 Gln Gly Ser Thr Tyr Cys Asp Ile Cys Lys Phe Gly Phe Glu Thr Pro  
 35 40 45  
 Glu Ser Ser Tyr Phe Ile Pro Gly Ala Thr Val Lys Leu Ser Cys Lys  
 50 55 60  
 Asp Arg Lys Thr Met Glu Glu Val Tyr Thr Asp Lys Ala Val Ser Asp  
 65 70 75 80  
 Lys Glu Gly Lys Tyr Lys Phe Ile Val His Asp Asp His Arg Asp Gln  
 85 90 95  
 Met Cys Asp Val Leu Leu Val Lys Ser Ser Asp Lys Thr Cys Ser Lys

Ile Ser Val Gly Arg Glu Lys Ser Arg Val Ile Leu Asn His Tyr Ser  
110  
115  
120  
125  
Gly Ile Ala Ser Gln Ile Arg His Ala Asn Asn Met Gly Phe Glu Lys  
130  
135  
140  
145  
Glu Val Ser Asp Val Phe Cys Ser Ala Leu Phe Gln Lys Tyr Met Val  
150  
155  
160  
Asp Glu Asp Glu Asp Asp Ile Lys Asn His Leu  
165  
170

<210> 299  
<211> 166  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..166:Ceres Seq. ID 1573608

Mot Leu Leu Val Leu Cys Ile Leu Pro Ala Ile Ala Met Ala Ala Arg  
1  
5  
10  
15  
Arg Ser Asn Ile Gly Lys Asn Thr Met Val Val Gln Gly Ser Thr Tyr  
20  
25  
30  
Cys Asp Ile Cys Lys Phe Gly Phe Glu Thr Pro Glu Ser Ser Tyr Phe  
35  
40  
45  
Ile Pro Gly Ala Thr Val Lys Leu Ser Cys Lys Asp Arg Lys Thr Met  
50  
55  
60  
Glu Glu Val Tyr Thr Asp Lys Ala Val Ser Asp Lys Glu Gly Lys Tyr  
65  
70  
75  
80  
Lys Phe Ile Val His Asp Asp His Arg Asp Gln Met Cys Asp Val Leu  
90  
95  
Leu Val Lys Ser Ser Asp Lys Thr Cys Ser Lys Ile Ser Val Gly Arg  
100  
105  
110  
Glu Lys Ser Arg Val Ile Leu Asn His Tyr Ser Gly Ile Ala Ser Gln  
115  
120  
125  
Ile Arg His Ala Asn Asn Met Gly Phe Glu Lys Val Ser Asp Val  
130  
135  
140  
Phe Cys Ser Ala Leu Phe Gln Lys Tyr Met Val Asp Glu Asp  
145  
150  
155  
Asp Ile Lys Asn His Leu  
160  
165

<210> 300  
<211> 154  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..154:Ceres Seq. ID 1573609

Mot Ala Ala Arg Arg Ser Asn Ile Gly Lys Asn Thr Met Val Val Gln  
1  
5  
10  
15  
Gly Ser Thr Tyr Cys Asp Ile Cys Lys Phe Gly Phe Glu Thr Pro Glu  
20  
25  
30  
Ser Ser Tyr Phe Ile Pro Gly Ala Thr Val Lys Leu Ser Cys Lys Asp  
35  
40  
45

Arg Lys Thr Met Glu Glu Val Tyr Thr Asp Lys Ala Val Ser Asp Lys  
50  
55  
60  
Glu Gly Lys Tyr Lys Phe Ile Val His Asp Asp His Arg Asp Gln Met  
65  
70  
75  
80  
Cys Asp Val Leu Leu Val Lys Ser Ser Asp Lys Thr Cys Ser Lys Ile  
85  
90  
95  
Ser Val Gly Arg Glu Lys Ser Arg Val Ile Leu Asn His Tyr Ser Gly  
100  
105  
110  
Ile Ala Ser Gln Ile Arg His Ala Asn Asn Met Gly Phe Glu Lys Glu  
115  
120  
125  
Val Ser Asp Val Phe Cys Ser Ala Leu Phe Gln Lys Tyr Met Val Asp  
130  
135  
140  
Glu Asp Glu Asp Asp Ile Lys Asn His Leu  
145  
150

<210> 301  
<211> 704  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..704:Ceres Seq. ID 1573861

<400>301  
aaaaacga atcaaatc ctacagttc aattttttg ttctgtttt tttttttt  
120  
125  
130  
135  
140  
145  
150  
155  
160  
165  
170  
175  
180  
185  
190  
195  
200  
205  
210  
215  
220  
225  
230  
235  
240  
245  
250  
255  
260  
265  
270  
275  
280  
285  
290  
295  
300  
305  
310  
315  
320  
325  
330  
335  
340  
345  
350  
355  
360  
365  
370  
375  
380  
385  
390  
395  
400  
405  
410  
415  
420  
425  
430  
435  
440  
445  
450  
455  
460  
465  
470  
475  
480  
485  
490  
495  
500  
505  
510  
515  
520  
525  
530  
535  
540  
545  
550  
555  
560  
565  
570  
575  
580  
585  
590  
595  
600  
605  
610  
615  
620  
625  
630  
635  
640  
645  
650  
655  
660  
665  
670  
675  
680  
685  
690  
695  
700  
705  
710  
715  
720  
725  
730  
735  
740  
745  
750  
755  
760  
765  
770  
775  
780  
785  
790  
795  
800  
805  
810  
815  
820  
825  
830  
835  
840  
845  
850  
855  
860  
865  
870  
875  
880  
885  
890  
895  
900  
905  
910  
915  
920  
925  
930  
935  
940  
945  
950  
955  
960  
965  
970  
975  
980  
985  
990  
995  
1000  
1005  
1010  
1015  
1020  
1025  
1030  
1035  
1040  
1045  
1050  
1055  
1060  
1065  
1070  
1075  
1080  
1085  
1090  
1095  
1100  
1105  
1110  
1115  
1120  
1125  
1130  
1135  
1140  
1145  
1150  
1155  
1160  
1165  
1170  
1175  
1180  
1185  
1190  
1195  
1200  
1205  
1210  
1215  
1220  
1225  
1230  
1235  
1240  
1245  
1250  
1255  
1260  
1265  
1270  
1275  
1280  
1285  
1290  
1295  
1300  
1305  
1310  
1315  
1320  
1325  
1330  
1335  
1340  
1345  
1350  
1355  
1360  
1365  
1370  
1375  
1380  
1385  
1390  
1395  
1400  
1405  
1410  
1415  
1420  
1425  
1430  
1435  
1440  
1445  
1450  
1455  
1460  
1465  
1470  
1475  
1480  
1485  
1490  
1495  
1500  
1505  
1510  
1515  
1520  
1525  
1530  
1535  
1540  
1545  
1550  
1555  
1560  
1565  
1570  
1575  
1580  
1585  
1590  
1595  
1600  
1605  
1610  
1615  
1620  
1625  
1630  
1635  
1640  
1645  
1650  
1655  
1660  
1665  
1670  
1675  
1680  
1685  
1690  
1695  
1700  
1705  
1710  
1715  
1720  
1725  
1730  
1735  
1740  
1745  
1750  
1755  
1760  
1765  
1770  
1775  
1780  
1785  
1790  
1795  
1800  
1805  
1810  
1815  
1820  
1825  
1830  
1835  
1840  
1845  
1850  
1855  
1860  
1865  
1870  
1875  
1880  
1885  
1890  
1895  
1900  
1905  
1910  
1915  
1920  
1925  
1930  
1935  
1940  
1945  
1950  
1955  
1960  
1965  
1970  
1975  
1980  
1985  
1990  
1995  
2000  
2005  
2010  
2015  
2020  
2025  
2030  
2035  
2040  
2045  
2050  
2055  
2060  
2065  
2070  
2075  
2080  
2085  
2090  
2095  
2100  
2105  
2110  
2115  
2120  
2125  
2130  
2135  
2140  
2145  
2150  
2155  
2160  
2165  
2170  
2175  
2180  
2185  
2190  
2195  
2200  
2205  
2210  
2215  
2220  
2225  
2230  
2235  
2240  
2245  
2250  
2255  
2260  
2265  
2270  
2275  
2280  
2285  
2290  
2295  
2300  
2305  
2310  
2315  
2320  
2325  
2330  
2335  
2340  
2345  
2350  
2355  
2360  
2365  
2370  
2375  
2380  
2385  
2390  
2395  
2400  
2405  
2410  
2415  
2420  
2425  
2430  
2435  
2440  
2445  
2450  
2455  
2460  
2465  
2470  
2475  
2480  
2485  
2490  
2495  
2500  
2505  
2510  
2515  
2520  
2525  
2530  
2535  
2540  
2545  
2550  
2555  
2560  
2565  
2570  
2575  
2580  
2585  
2590  
2595  
2600  
2605  
2610  
2615  
2620  
2625  
2630  
2635  
2640  
2645  
2650  
2655  
2660  
2665  
2670  
2675  
2680  
2685  
2690  
2695  
2700  
2705  
2710  
2715  
2720  
2725  
2730  
2735  
2740  
2745  
2750  
2755  
2760  
2765  
2770  
2775  
2780  
2785  
2790  
2795  
2800  
2805  
2810  
2815  
2820  
2825  
2830  
2835  
2840  
2845  
2850  
2855  
2860  
2865  
2870  
2875  
2880  
2885  
2890  
2895  
2900  
2905  
2910  
2915  
2920  
2925  
2930  
2935  
2940  
2945  
2950  
2955  
2960  
2965  
2970  
2975  
2980  
2985  
2990  
2995  
3000  
3005  
3010  
3015  
3020  
3025  
3030  
3035  
3040  
3045  
3050  
3055  
3060  
3065  
3070  
3075  
3080  
3085  
3090  
3095  
3100  
3105  
3110  
3115  
3120  
3125  
3130  
3135  
3140  
3145  
3150  
3155  
3160  
3165  
3170  
3175  
3180  
3185  
3190  
3195  
3200  
3205  
3210  
3215  
3220  
3225  
3230  
3235  
3240  
3245  
3250  
3255  
3260  
3265  
3270  
3275  
3280  
3285  
3290  
3295  
3300  
3305  
3310  
3315  
3320  
3325  
3330  
3335  
3340  
3345  
3350  
3355  
3360  
3365  
3370  
3375  
3380  
3385  
3390  
3395  
3400  
3405  
3410  
3415  
3420  
3425  
3430  
3435  
3440  
3445  
3450  
3455  
3460  
3465  
3470  
3475  
3480  
3485  
3490  
3495  
3500  
3505  
3510  
3515  
3520  
3525  
3530  
3535  
3540  
3545  
3550  
3555  
3560  
3565  
3570  
3575  
3580  
3585  
3590  
3595  
3600  
3605  
3610  
3615  
3620  
3625  
3630  
3635  
3640  
3645  
3650  
3655  
3660  
3665  
3670  
3675  
3680  
3685  
3690  
3695  
3700  
3705  
3710  
3715  
3720  
3725  
3730  
3735  
3740  
3745  
3750  
3755  
3760  
3765  
3770  
3775  
3780  
3785  
3790  
3795  
3800  
3805  
3810  
3815  
3820  
3825  
3830  
3835  
3840  
3845  
3850  
3855  
3860  
3865  
3870  
3875  
3880  
3885  
3890  
3895  
3900  
3905  
3910  
3915  
3920  
3925  
3930  
3935  
3940  
3945  
3950  
3955  
3960  
3965  
3970  
3975  
3980  
3985  
3990  
3995  
4000  
4005  
4010  
4015  
4020  
4025  
4030  
4035  
4040  
4045  
4050  
4055  
4060  
4065  
4070  
4075  
4080  
4085  
4090  
4095  
4100  
4105  
4110  
4115  
4120  
4125  
4130  
4135  
4140  
4145  
4150  
4155  
4160  
4165  
4170  
4175  
4180  
4185  
4190  
4195  
4200  
4205  
4210  
4215  
4220  
4225  
4230  
4235  
4240  
4245  
4250  
4255  
4260  
4265  
4270  
4275  
4280  
4285  
4290  
4295  
4300  
4305  
4310  
4315  
4320  
4325  
4330  
4335  
4340  
4345  
4350  
4355  
4360  
4365  
4370  
4375  
4380  
4385  
4390  
4395  
4400  
4405  
4410  
4415  
4420  
4425  
4430  
4435  
4440  
4445  
4450  
4455  
4460  
4465  
4470  
4475  
4480  
4485  
4490  
4495  
4500  
4505  
4510  
4515  
4520  
4525  
4530  
4535  
4540  
4545  
4550  
4555  
4560  
4565  
4570  
4575  
4580  
4585  
4590  
4595  
4600  
4605  
4610  
4615  
4620  
4625  
4630  
4635  
4640  
4645  
4650  
4655  
4660  
4665  
4670  
4675  
4680  
4685  
4690  
4695  
4700  
4705  
4710  
4715  
4720  
4725  
4730  
4735  
4740  
4745  
4750  
4755  
4760  
4765  
4770  
4775  
4780  
4785  
4790  
4795  
4800  
4805  
4810  
4815  
4820  
4825  
4830  
4835  
4840  
4845  
4850  
4855  
4860  
4865  
4870  
4875  
4880  
4885  
4890  
4895  
4900  
4905  
4910  
4915  
4920  
4925  
4930  
4935  
4940  
4945  
4950  
4955  
4960  
4965  
4970  
4975  
4980  
4985  
4990  
4995  
5000  
5005  
5010  
5015  
5020  
5025  
5030  
5035  
5040  
5045  
5050  
5055  
5060  
5065  
5070  
5075  
5080  
5085  
5090  
5095  
5100  
5105  
5110  
5115  
5120  
5125  
5130  
5135  
5140  
5145  
5150  
5155  
5160  
5165  
5170  
5175  
5180  
5185  
5190  
5195  
5200  
5205  
5210  
5215  
5220  
5225  
5230  
5235  
5240  
5245  
5250  
5255  
5260  
5265  
5270  
5275  
5280  
5285  
5290  
5295  
5300  
5305  
5310  
5315  
5320  
5325  
5330  
5335  
5340  
5345  
5350  
5355  
5360  
5365  
5370  
5375  
5380  
5385  
5390  
5395  
5400  
5405  
5410  
5415  
5420  
5425  
5430  
5435  
5440  
5445  
5450  
5455  
5460  
5465  
5470  
5475  
5480  
5485  
5490  
5495  
5500  
5505  
5510  
5515  
5520  
5525  
5530  
5535  
5540  
5545  
5550  
5555  
5560  
5565  
5570  
5575  
5580  
5585  
5590  
5595  
5600  
5605  
5610  
5615  
5620  
5625  
5630  
5635  
5640  
5645  
5650  
5655  
5660  
5665  
5670  
5675  
5680  
5685  
5690  
5695  
5700  
5705  
5710  
5715  
5720  
5725  
5730  
5735  
5740  
5745  
5750  
5755  
5760  
5765  
5770  
5775  
5780  
5785  
5790  
5795  
5800  
5805  
5810  
5815  
5820  
5825  
5830  
5835  
5840  
5845  
5850  
5855  
5860  
5865  
5870  
5875  
5880  
5885  
5890  
5895  
5900  
5905  
5910  
5915  
5920  
5925  
5930  
5935  
5940  
5945  
5950  
5955  
5960  
5965  
5970  
5975  
5980  
5985  
5990  
5995  
6000  
6005  
6010  
6015  
6020  
6025  
6030  
6035  
6040  
6045  
6050  
6055  
6060  
6065  
6070  
6075  
6080  
6085  
6090  
6095  
6100  
6105  
6110  
6115  
6120  
6125  
6130  
6135  
6140  
6145  
6150  
6155  
6160  
6165  
6170  
6175  
6180  
6185  
6190  
6195  
6200  
6205  
6210  
6215  
6220  
6225  
6230  
6235  
6240  
6245  
6250  
6255  
6260  
6265  
6270  
6275  
6280  
6285  
6290  
6295  
6300  
6305  
6310  
6315  
6320  
6325  
6330  
6335  
6340  
6345  
6350  
6355  
6360  
6365  
6370  
6375  
6380  
6385  
6390  
6395  
6400  
6405  
6410  
6415  
6420  
6425  
6430  
6435  
6440  
6445  
6450  
6455  
6460  
6465  
6470  
6475  
6480  
6485  
6490  
6495  
6500  
6505  
6510  
6515  
6520  
6525  
6530  
6535  
6540  
6545  
6550  
6555  
6560  
6565  
6570  
6575  
6580  
6585  
6590  
6595  
6600  
6605  
6610  
6615  
6620  
6625  
6630  
6635  
6640  
6645  
6650  
6655  
6660  
6665  
6670  
6675  
6680  
6685  
6690  
6695  
6700  
6705  
6710  
6715  
6720  
6725  
6730  
6735  
6740  
6745  
6750  
6755  
6760  
6765  
6770  
6775  
6780  
6785  
6790  
6795  
6800  
6805  
6810  
6815  
6820  
6825  
6830  
6835  
6840  
6845  
6850  
6855  
6860  
6865  
6870  
6875  
6880  
6885  
6890  
6895  
6900  
6905  
6910  
6915  
6920  
6925  
6930  
6935  
6940  
6945  
6950  
6955  
6960  
6965  
6970  
6975  
6980  
6985  
6990  
6995  
7000  
7005  
7010  
7015  
7020  
7025  
7030  
7035  
7040  
7045  
7050  
7055  
7060  
7065  
7070  
7075  
7080  
7085  
7090  
7095  
7100  
7105  
7110  
7115  
7120  
7125  
7130  
7135  
7140  
7145  
7150  
7155  
7160  
7165  
7170  
7175  
7180  
7185  
7190  
7195  
7200  
7205  
7210  
7215  
7220  
7225  
7230  
7235  
7240  
7245  
7250  
7255  
7260  
7265  
7270  
7275  
7280  
7285  
7290  
7295  
7300  
7305  
7310  
7315  
7320  
7325  
7330  
7335  
7340  
7345  
7350  
7355  
7360  
7365  
7370  
7375  
7380  
7385  
7390  
7395  
7400  
7405  
7410  
7415  
7420  
7425  
7430  
7435  
7440  
7445  
7450  
7455  
7460  
7465  
7470  
7475  
7480  
7485  
7490  
7495  
7500  
7505  
7510  
7515  
7520  
7525  
7530  
7535  
7540  
7545  
7550  
7555  
7560  
7565  
7570  
7575  
7580  
7585  
7590  
7595  
7600  
7605  
7610  
7615  
7620  
7625  
7630  
7635  
7640  
7645  
7650  
7655  
7660  
7665  
7670  
7675  
7680  
7685  
7690  
7695  
7700  
7705  
7710  
7715  
7720  
7725  
7730  
7735  
7740  
7745  
7750  
7755  
7760  
7765  
7770  
7775  
7780  
7785  
7790  
7795  
7800  
7805  
7810  
7815  
7820  
7825  
7830  
7835  
7840  
7845  
7850  
7855  
7860  
7865  
7870  
7875  
7880  
7885  
7890  
7895  
7900  
7905  
7910  
7915  
7920  
7925  
7930  
7935  
7940  
7945  
7950  
7955  
7960  
7965  
7970  
7975  
7980  
7985  
7990  
7995  
8000  
8005  
8010  
8015  
8020  
8025  
8030  
8035  
8040  
8045  
8050  
8055  
8060  
8065  
8070  
8075  
8080  
8085  
8090  
8095  
8100  
8105  
8110  
8115  
8120  
8125  
8130  
8135  
8140  
8145  
8150  
8155  
8160  
8165  
8170  
8175  
8180  
8185  
8190  
8195  
8200  
8205  
8210  
8215  
8220  
8225  
8230  
8235  
8240  
8245  
8250  
8255  
8260  
8265  
8270  
8275  
8280  
8285  
8290  
8295  
8300  
8305  
8310  
8315  
8320  
8325  
8330  
8335  
8340  
8345  
8350  
8355  
8360  
8365  
8370  
8375  
8380  
8385  
8390  
8395  
8400  
8405  
8410  
8415  
8420  
8425  
8430  
8435  
8440  
8445  
8450  
8455  
8460  
8465  
8470  
8475  
8480  
8485  
8490  
8495  
8500  
8505  
8510  
8515  
8520  
8525  
8530  
8535  
8540  
8545  
8550  
8555  
8560  
8565  
8570  
8575  
8580  
8585  
8590  
8595  
8600  
8605  
8610  
8615  
8620  
8625  
8630  
8635  
8640  
8645  
8650  
8655  
8660  
8665  
8670  
8675  
8680  
8685  
8690  
8695  
8700  
8705  
8710  
8715  
8720  
8725  
8730  
8735  
8740  
8745  
8750  
8755  
8760  
8765  
8770  
8775  
8780  
8785  
8790  
8795  
8800  
8805  
8810  
8815  
8820  
8825  
8830  
8835  
8840  
8845  
8850  
8855  
8860  
8865  
8870  
8875  
8880  
8885  
8890  
8895  
8900  
8905  
8910

85 90 95  
 Leu Trp Ala Ser Leu Asn Leu Ile Thr Ala Ile Asp Lys Leu Pro Val  
 100 105 110  
 Ile Ser Gly Phe Glu Leu Val Gly Ile Leu Phe Ser Thr Trp Phe  
 115 120 125  
 Thr Tyr Arg Tyr Leu Leu Phe Lys Lys Pro Asp Arg Gln Glu Leu Ser Lys  
 130 135 140  
 Ile Val Lys Lys Ser Val Ala Asp Ile Leu Gly Thr Val Asn Leu Val  
 145 150 155 160  
 Cys Val Ile Ile Leu His Leu Trp Lys Met Ile Cys Leu Gln Val Cys  
 165 170 175  
 Lys Ile Thr

&lt;210&gt; 303

&lt;211&gt; 127

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..127:Ceres Seq. ID 1573863

&lt;400&gt;303

Met Val Lys Ala Ser Gly Glu Ser Ser Asp Ser Ser Thr Asp Leu Asp  
 1 5 10 15  
 Val Val Ser Thr Ile Gln Asn Val Trp Asp Lys Ser Glu Asn Arg Leu  
 20 25 30  
 Gly Leu Ile Gly Leu Gly Phe Ala Gly Ile Val Ala Leu Trp Ala Ser  
 35 40 45  
 Leu Asn Leu Ile Thr Ala Ile Asp Lys Leu Pro Val Ile Ser Ser Gly  
 50 55 60  
 Phe Glu Leu Val Gly Ile Leu Phe Ser Thr Trp Phe Thr Tyr Arg Tyr  
 65 70 75 80  
 Leu Leu Phe Lys Pro Asp Arg Gln Glu Leu Ser Lys Ile Val Lys Lys  
 85 90 95  
 Ser Val Ala Asp Ile Leu Gly Thr Val Asn Leu Val Cys Val Ile Ile  
 100 105 110  
 Leu His Leu Trp Lys Met Ile Cys Leu Gln Val Cys Lys Ile Thr  
 115 120 125

&lt;210&gt; 304

&lt;211&gt; 662

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..662:Ceres Seq. ID 1574093

&lt;400&gt;304

attttttt tctctctc ctccctaagc aaactaaaa caagctatgg ctggtatgct  
 60  
 tcccgagtt gaggtagc aaagggggc ctccaaggt ggtgtctctc cagttgaatc  
 120  
 ctgaacaca gctctgtgg cggctggcg gggacagtc tggacacggc gaccatcgtt  
 180  
 ctctctttac attcttaac atggagcca ccagggccat gctctctct cggagagaag  
 240  
 tttagaat aaattctatg gagaagaca cagtagagaa cttagcgag cagccaaga  
 300  
 ggaacagc aggtttaaca agcgttgag aatccacca cgtacaagt caggcaaat  
 360  
 ggaagaca aggggaata attggagca ggaagggtt aacctctgg ggaattacc  
 420  
 accgagttg ctggggttaa agagagccg aggaagttg atggaatgt tcaagcgcg  
 480

agtaggaa caacaagatt gtgctatag tctagaccg tccaagaag gtgagacct  
 540  
 ggtacacta catgtgcc ataatgtca ctccatagc ttatgctt ggtagaac  
 600  
 taatgttat tgccatatt gtagaactg tatttgaat taaatgtat attttgatg  
 660  
 tc

&lt;210&gt; 305

&lt;211&gt; 198

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..198:Ceres Seq. ID 1574094

&lt;400&gt;305

Met Ala Gly Met Leu Pro Gly Val Glu Cys Ala Arg Arg Arg Arg Phe  
 1 5 10 15  
 His Gly Gly Ala Pro Pro Ile Glu Ser Ser Asn Thr Ala Ser Val Ala  
 20 25 30  
 Ala Ala Ala Gly His Val Trp Thr Arg Arg Pro Ser Phe Ser Leu Tyr  
 35 40 45  
 Thr Thr Asn His Glu Ser His Gln Ala His Val Ser Phe Ser Glu Arg  
 50 55 60  
 Ser Val Arg Asn Lys Ser Tyr Gly Glu Asp Asn Asp Glu Lys Leu Asp  
 65 70 75 80  
 Gly Ala Ala Lys Glu Ala Lys Gln Arg Leu Asn Lys Arg Leu Arg Ile  
 85 90 95  
 Pro Pro Arg Thr Ser Ser Gly Lys Met Val Lys Thr Lys Gly Ile Asn  
 100 105 110  
 Trp Ser Lys Glu Arg Val Asn Leu Ser Gly Thr Tyr Arg Pro Arg Trp  
 115 120 125  
 Val Gly Leu Lys Lys Ser Arg Gly Arg Leu Met Glu Trp Phe Lys Arg  
 130 135 140  
 Arg Val Arg Glu Gln Gln Asp Cys Ala Ile Cys Leu Asp Arg Phe Lys  
 145 150 155 160  
 Lys Gly Glu Thr Leu Val His Leu Pro Cys Ala His Lys Phe His Ser  
 165 170 175  
 Ile Cys Leu Leu Pro Trp Leu Asp Thr Asn Val Tyr Cys Pro Tyr Cys  
 180 185 190  
 Arg Thr Asp Ile Trp Asn  
 195

&lt;210&gt; 306

&lt;211&gt; 195

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..195:Ceres Seq. ID 1574095

&lt;400&gt;306

Met Leu Pro Gly Val Glu Cys Ala Arg Arg Arg Arg Phe His Gly Gly  
 1 5 10 15  
 Ala Pro Pro Ile Glu Ser Ser Asn Thr Ala Ser Val Ala Ala Ala  
 20 25 30  
 Gly His Val Trp Thr Arg Arg Pro Ser Phe Ser Leu Tyr Thr Asn  
 35 40 45  
 His Glu Ser His Gln Ala His Val Ser Phe Ser Glu Arg Ser Val Arg

50 55 60  
 Asn Lys Ser Tyr Gly Glu Asp Asn Asp Glu Lys Leu Asp Gly Ala Ala  
 65 70 75 80  
 Lys Glu Ala Lys Glu Arg Leu Asn Lys Arg Leu Arg Ile Pro Pro Arg  
 85 90 95  
 Thr Ser Ser Gly Lys Met Val Lys Thr Lys Gly Ile Asn Trp Ser Lys  
 100 105 110  
 Glu Arg Val Asn Leu Ser Gly Thr Tyr Arg Pro Arg Trp Val Gly Leu  
 115 120 125  
 Lys Lys Ser Arg Gly Arg Leu Met Glu Trp Phe Lys Arg Arg Val Arg  
 130 135 140  
 Glu Gln Gln Asp Cys Ala Ile Cys Leu Asp Arg Phe Lys Lys Gly Glu  
 145 150 155 160  
 Thr Leu Val His Leu Pro Cys Ala His Lys Phe His Ser Ile Cys Leu  
 165 170 175  
 Leu Pro Trp Leu Asp Thr Asn Val Tyr Cys Pro Tyr Cys Arg Thr Asp  
 180 185 190  
 Ile Trp Asn  
 195

&lt;210&gt; 307

&lt;211&gt; 725

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..725:Ceres Seq. ID 1580388

&lt;400&gt;307

ctcacacaa aatacacaaa cttagatcag tctcaaaagg ggaacacaaa cttaaagaa  
 acataaagg gcaacacaaa tcaacaaa gatcaaatg agactaaga gaagggcaaa  
 aagtggaag caatggctac caatgggaag gctctcttg tatgtcatt gtaggttt  
 ctgtgtatg ctgtgtgac tccatggcg ggaacccat tcaagaaag cgtagtctc  
 ggaggeagt cagcgttcc aaacttcgg accaacagg aaattcaaa acttggaagg  
 tactgtggt agaaatcaa tcaacaaga cagaacaggc aggaacacat aggtaccat  
 gcgaacag acacggcat tccgaatca tgcgaattg gccgtagt gtcgtcag  
 aaacagtgc tgcgtgact caaatctat ctgaagattg agtcactca acccaatgc  
 tctaccaga tgttgact tgttgttt attcaacat gccctcattc taagcagttg  
 ctccgtttca ctccgtgt cagtcctgc tactaaett attctctt attgactta  
 aatttcata atatgataa gaaagacta aaagtgat gatacaagc tattaaagat  
 gggtaaatg ttggtttca tgatatgtt acgtgttca taataaaaa caagttgta  
 ttagg

&lt;210&gt; 308

&lt;211&gt; 147

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..147:Ceres Seq. ID 1580389

&lt;400&gt;308

Met Ala Thr Met Leu Lys Val Ser Leu Val Leu Ser Leu Leu Gly Phe  
 1 5 10 15  
 Leu Val Ile Ala Val Val Thr Pro Ser Ala Ala Asn Pro Phe Arg Lys  
 20 25 30  
 Ser Val Val Leu Gly Gly Lys Ser Gly Val Pro Asn Ile Arg Thr Asn

35 40 45  
 Arg Glu Ile Gln Gln Leu Gly Arg Tyr Cys Val Glu Gln Phe Asn Gln  
 50 55 60  
 Gln Ala Gln Asn Glu Gln Gly Asn Ile Gly Ser Ile Ala Lys Thr Asp  
 65 70 75 80  
 Thr Ala Ile Ser Asn Pro Leu Gln Phe Ser Arg Val Val Ser Ala Gln  
 85 90 95  
 Lys Gln Val Val Ala Gly Leu Lys Tyr Tyr Leu Arg Ile Glu Val Thr  
 100 105 110  
 Gln Pro Asn Gly Ser Thr Arg Met Phe Asp Ser Val Val Ile Gln  
 115 120 125  
 Pro Trp Leu His Ser Lys Gln Leu Leu Gly Phe Thr Pro Val Val Ser  
 130 135 140  
 Pro Val Tyr  
 145

&lt;210&gt; 309

&lt;211&gt; 144

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..144:Ceres Seq. ID 1580390

&lt;400&gt;309

Met Leu Lys Val Ser Leu Val Leu Ser Leu Leu Gly Phe Leu Val Ile  
 1 5 10 15  
 Ala Val Val Thr Pro Ser Ala Ala Asn Pro Phe Arg Lys Ser Val Val  
 20 25 30  
 Leu Gly Gly Lys Ser Gly Val Pro Asn Ile Arg Thr Asn Arg Glu Ile  
 35 40 45  
 Gln Gln Leu Gly Arg Tyr Cys Val Glu Gln Phe Asn Gln Gln Ala Gln  
 50 55 60  
 Asn Glu Gln Gly Asn Ile Gly Ser Ile Ala Lys Thr Asp Thr Ala Ile  
 65 70 75 80  
 Ser Asn Pro Leu Gln Phe Ser Arg Val Val Ser Ala Gln Lys Gln Val  
 85 90 95  
 Val Ala Gly Leu Lys Tyr Tyr Leu Arg Ile Glu Val Thr Gln Pro Asn  
 100 105 110  
 Gly Ser Thr Arg Met Phe Asp Ser Val Val Ile Gln Pro Trp Leu  
 115 120 125  
 His Ser Lys Gln Leu Leu Gly Phe Thr Pro Val Val Ser Pro Val Tyr  
 130 135 140

&lt;210&gt; 310

&lt;211&gt; 675

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..675:Ceres Seq. ID 1582959

&lt;400&gt;310

tgggaaggc aacaaaceta atcaacttan atcttatcta cttctattt ctttttacc  
 60  
 aaattacccg tcttaacta tgggaaggc gttttccat atctcttgg cttttgccc  
 120  
 agctccagct ttacgttgg gcgcaagaa ggtccagta ggtctcttg accaaagaa  
 180  
 ctaccctoga tatgtggcg gatattccgg cgttgagac aatgggttac ccttggggg  
 240

cgtcgggga ggtgtgtctg gtccggagg taacttggt tatggggat ttgggtgctc 300  
 tgggcggc ttggcggtg gttgggggg tggagcaggc agtgattag ggggtgctt 360  
 agtggggga agtgggattg gtccggaac cagtggagg agtaccggag ggttcatt 420  
 cccgttggt gtacttgg ttttaaggc gtacacggt cctattaag ctggtctag 480  
 ctaagatga tgcataata ataattatc atactcttt aggttttaa acttggat 540  
 ctgaattat cttgagtgt ttaactgag tettaagta ctatttaac gtaggttga 600  
 atcagctag tggctgtgct gtcttggtt tggctcatt tcaattcta gttgacctt 660  
 ttgaggtt tcacc 675

&lt;210&gt; 311

&lt;211&gt; 115

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..115:Ceres Seq. ID 1582960

&lt;400&gt;311

Met Ala Lys Trp Phe Thr Ile Phe Leu Val Phe Ala Leu Ala Ser 1  
 1 5 10 15  
 Ala Leu Ala Cys Gly Ala Arg Asn Val Pro Val Gly Leu Ser Asp Gln 20  
 25 30  
 Lys Asn Tyr Leu Gly Tyr Gly Gly Tyr Ser Gly Val Gly Asp Asn 35  
 40 45  
 Gly Leu Pro Phe Gly Gly Val Gly Gly Val Ser Gly Pro Gly Gly 50  
 55 60  
 Asn Leu Gly Tyr Gly Gly Phe Gly Gly Ala Gly Gly Leu Gly Gly 65  
 70 75 80  
 Gly Leu Gly Gly Ala Gly Ser Gly Leu Gly Gly Gly Leu Gly Gly 85  
 90 95  
 Gly Ser Gly Ile Gly Ala Gly Thr Ser Gly Gly Ser Thr Gly Gly Val 100  
 105 110  
 His Pro 115

&lt;210&gt; 312

&lt;211&gt; 1076

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..1076:Ceres Seq. ID 1663221

<400>312  
 aacagagct taccgattct ttctgtctca acatttaggg ttccagagat cggctgattt 60  
 tcatctac agtatggcg actcaatca gcaagaag aaagtgtga gggatggg 120  
 tatctcagc tgaattgat aggttctca caagagact agcagagat ggttacttg 180  
 ggttgaggt taggttact ccaatgaga ctgagattat catcagagat actgtactc 240  
 agaatgttt tggtagaag gggagagaa ttagggaatt gacatcttt gtccagaaga 300  
 gattcaagtt tccagttgac agtttgagc tccatgtga gaaggttaac aacagaggtc 360  
 ttgtgcat tctcagct ggtctctac gttacaagct tctcgtggt ctgtgttc 420  
 gcaaggcctg ttatgggtt ttggatttg ttatggag tggagctag ggtatgag 480  
 tcatgtgag tgaagaact cgtgtgcac gtgtcagtc catgaagtc aaggtggt 540  
 acatggttc atctggkca ccaactaag aatacatga tgcagctg agcatgttt 600  
 tctcagca ggggtgttg ggaatcaag tgaatcat cttgtactg gacctcagg 660  
 gcaatcagg accaagaca ccaatgctg atgttgtat cattcatgt cctaaagatg 720  
 atgttgtcta ctctgcacct gctcaggctg ctgtccagt tactttgtg caagaagctc 780

cactacaac cgtagattac cctgagatga ttctccagt ggcctagaga agaccitttt 840  
 tactattact caatggatt ttgtctttt tgtataact ttctacttt tgaactctc 900  
 tagctatc tctataacc tccaagaga caagttttt cctagttgt tcttaacct 960  
 atgtcaagt ttgtgaggt tgaatggtt tagtccatg atttttgt tcttaagact 1020  
 cgtataaga aaaagaatga cgaacaact gccgggtatt tgatgttgtt actttg 1076

&lt;210&gt; 313

&lt;211&gt; 250

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..250:Ceres Seq. ID 1663222

&lt;400&gt;313

Met Ala Thr Gln Ile Ser Lys Lys Arg Lys Phe Val Ala Asp Gly Val 1  
 5 10 15  
 Phe Tyr Ala Glu Leu Asn Glu Val Leu Thr Arg Glu Leu Ala Glu Asp 20  
 25 30  
 Gly Tyr Ser Gly Val Glu Val Arg Val Thr Pro Met Arg Thr Glu Ile 35  
 40 45  
 Ile Ile Arg Ala Thr Arg Thr Gln Asn Val Leu Gly Glu Lys Gly Arg 50  
 55 60  
 Arg Ile Arg Glu Leu Thr Ser Leu Val Gln Lys Arg Phe Lys Phe Pro 65  
 70 75 80  
 Val Asp Ser Val Glu Leu Tyr Ala Glu Lys Val Asn Asn Arg Gly Leu 85  
 90 95  
 Cys Ala Ile Ala Gln Ala Glu Ser Leu Arg Tyr Lys Leu Leu Gly Gly 100  
 105 110  
 Leu Ala Val Arg Arg Ala Cys Tyr Gly Val Leu Arg Phe Val Met Glu 115  
 120 125  
 Ser Gly Ala Lys Gly Cys Glu Val Ile Val Ser Gly Lys Leu Arg Ala 130  
 135 140  
 Ala Arg Ala Lys Ser Met Lys Phe Lys Asp Gly Tyr Met Val Ser Ser 145  
 150 155 160  
 Gly Gln Pro Thr Lys Glu Tyr Ile Asp Ala Ala Val Arg His Val Leu 165  
 170 175  
 Leu Arg Gln Gly Val Leu Gly Ile Lys Val Lys Ile Met Leu Asp Trp 180  
 185 190  
 Asp Pro Thr Gly Lys Ser Gly Pro Lys Thr Pro Leu Pro Asp Val Val 195  
 200 205  
 Ile Ile His Ala Pro Lys Asp Asp Val Val Tyr Ser Ala Pro Ala Gln 210  
 215 220  
 Ala Ala Ala Pro Val Thr Leu Val Gln Glu Ala Pro Leu Thr Thr Val 225  
 230 235  
 Asp Tyr Pro Glu Met Ile Pro Pro Val Ala 240  
 245 250

&lt;210&gt; 314

&lt;211&gt; 207

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..207:Ceres Seq. ID 1663223

&lt;400&gt;314

Met Arg Thr Glu Ile Ile Arg Ala Thr Arg Thr Gln Asn Val Leu  
1 5 10 15  
Gly Glu Lys Gly Arg Arg Ile Arg Glu Leu Thr Ser Leu Val Gln Lys  
20 25 30  
Arg Phe Lys Phe Pro Val Asp Ser Val Glu Leu Tyr Ala Glu Lys Val  
35 40 45  
Asn Asn Arg Gly Leu Cys Ala Ile Ala Gln Ala Glu Ser Leu Arg Tyr  
50 55 60  
Lys Leu Leu Gly Gly Leu Ala Val Arg Arg Ala Cys Tyr Gly Val Leu  
65 70 75 80  
Arg Phe Val Met Glu Ser Gly Ala Lys Gly Cys Glu Val Ile Val Ser  
85 90 95  
Gly Lys Leu Arg Ala Ala Arg Ala Lys Ser Met Lys Phe Lys Asp Gly  
100 105 110  
Tyr Met Val Ser Ser Gly Gln Pro Thr Lys Glu Tyr Ile Asp Ala Ala  
115 120 125  
Val Arg His Val Leu Leu Arg Gln Gly Val Leu Gly Ile Lys Val Lys  
130 135 140  
Ile Met Leu Asp Trp Asp Pro Thr Gly Lys Ser Gly Pro Lys Thr Pro  
145 150 155  
Leu Pro Asp Val Val Ile Ile His Ala Pro Lys Asp Asp Val Val Tyr  
160 165 170 175  
Ser Ala Pro Ala Gln Ala Ala Pro Val Thr Leu Val Gln Glu Ala  
180 185  
Pro Leu Thr Thr Val Asp Tyr Pro Glu Met Ile Pro Pro Val Ala  
195 200 205

&lt;210&gt; 315

&lt;211&gt; 488

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..488:Ceres Seq. ID 1663275

&lt;400&gt;315

gtgactaaat ttctagggtt ttgaagagtc tctctgtcgc cogtttgtta ttctctgtg  
60  
agcaaatcga cgaatgggt cactcaatg tatggaactc tcatacgaag aagtaacgtc  
120  
ctgatatcgc ttatgcctg gtgtgggga actcgcaagg ttgtatcgg aagtatggt  
180  
tgaactgctg cagcactgt tctcgtagca acgcaagga gatggattc attaatgac  
240  
gttaatcag caccacactc atgattgatg cttaatgata taacatgaa ggcgtcgatg  
300  
ggataggctt ttaagccttt gtatgttttg aaattttac ttttgagaac catgtttatt  
360  
ttgggagtta attaagttgt tgaacctcat taagcatgdc ttattttgga ttaatgatgt  
420  
tttgctatt ctcgtatttt tgtttatca gtcaaatttg actaagaaga atttgcttgt  
488  
tattagcc

&lt;210&gt; 316

&lt;211&gt; 56

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..56:Ceres Seq. ID 1663276

&lt;400&gt;316

Met Gly His Ser Asn Val Trp Asn Ser His Pro Lys Lys Tyr Gly Pro  
1 5 10 15

Gly Ser Arg Leu Cys Arg Val Cys Gly Asn Ser His Gly Leu Ile Arg  
20 25 30  
Lys Tyr Gly Leu Asn Cys Cys Arg Gln Cys Phe Arg Ser Asn Ala Lys  
35 40 45  
Glu Ile Gly Phe Ile Lys Tyr Arg  
50 55

&lt;210&gt; 317

&lt;211&gt; 38

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..38:Ceres Seq. ID 1663277

&lt;400&gt;317

Met Glu Leu Ser Ser Glu Glu Val Arg Ser Trp Ile Ser Phe Met Pro  
1 5 10 15  
Cys Val Arg Glu Glu Ala Arg Ser Asp Pro Glu Val Trp Phe Glu Leu  
20 25 30  
Leu Gln Thr Val Phe Pro  
35

&lt;210&gt; 318

&lt;211&gt; 759

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..759:Ceres Seq. ID 1665304

&lt;400&gt;318

atcaaccaca acaaaacata aaaaacaagt ggaagettta aaacgagagg gagagagcaa  
60  
aaatgcgacg gtgggaacg tacgtgacgg aagttcogct aaaagatcg gcgagagaa  
120  
actacaagag gtgggaacg gagaacatg tcttcctga tgcatacgc caccacatcc  
180  
aaatgtttac ctttcacgaa ggcaacatg actctcacg gtctacagg agtggaaact  
240  
acacattgga tggaaaggag gagggtttca agggagagag agagatagac gatgagacca  
300  
aaacgttgac gttaaaggga cttgaggttc acgtgatgga gcagctcaaa gtgtacgacg  
360  
tgcttacc aattatccc aaattgagg atactgcgt cggcaaaatc actttaatat  
420  
gggagaacg caacatgat tcccagaac caaaggccta catgaaatc gtaagagact  
480  
tggttcgca catgggaac cagcttagca aaactaatc atcaatcca cagtcogct  
540  
cgtcgtcatc atcatcatca tcatacatc atcatcatc atcatcatc tcatacatc  
600  
tcactatc tcgatttata agttaagatg ttttcagtat aataaatggg gtctgtgga  
660  
tcgttcatt ctagtggaa acggttggt tctgtatgat gttcogatat attgttatgt  
720  
tcatacatc atgtcgggtt cgatataatg attcttaag  
759

&lt;210&gt; 319

&lt;211&gt; 151

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..151:Ceres Seq. ID 1665305

&lt;400&gt;319

Met Ala Thr Ser Gly Thr Tyr Val Thr Glu Val Pro Leu Lys Gly Ser  
1 5 10 15  
Ala Glu Lys Tyr Tyr Lys Arg Trp Lys Asn Glu Asn His Val Phe Pro  
20 25 30  
Asp Ala Ile Gly His His Ile Gln Asn Val Thr Val His Glu Gly Glu  
35 40 45  
His Asp Ser His Gly Ser Ile Arg Ser Trp Asn Tyr Thr Trp Asp Gly  
50 55 60  
Lys Glu Glu Val Phe Lys Glu Arg Arg Glu Ile Asp Asp Glu Thr Lys  
65 70 75 80  
Thr Leu Thr Leu Arg Gly Leu Glu Gly His Val Met Glu Gln Leu Lys  
85 90 95  
Val Tyr Asp Val Val Tyr Gln Phe Ile Pro Lys Ser Glu Asp Thr Cys  
100 105 110  
Ile Gly Lys Ile Thr Leu Ile Trp Glu Lys Arg Asn Asp Asp Ser Pro  
115 120 125  
Glu Pro Ser Gly Tyr Met Lys Phe Val Lys Ser Leu Val Ala Asp Met  
130 135 140  
Gly Asn His Val Ser Lys Thr  
145 150

&lt;210&gt; 320

&lt;211&gt; 602

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..602:Ceres Seq. ID 1709970

&lt;400&gt;320

gaagtcctt atataaagt atcagagaa ttagggttt ctagtctagt gtctcgga  
60  
gcagcagag cagagagag aagacaaga agaagagag cagctactac catgtcttg  
120  
ggacggagct tgcctgttc agtggcaga agattaccg aggaagagga ataggttta  
180  
tcgatctga ttctcaggtt ttctgtttc ttaactcaa atglaagagg tactccata  
240  
acaagtgaa gccatccaag ctgcatgga ctgccatga caaagagcaa cncagaagg  
300  
atcagacaa agagctcttg aagagaaga gactgccac caaagaagca tactcaaggt  
360  
ccattgttg tgcaccttg gaagtaattc agaagaagag agctgagaag cctgaagtc  
420  
gtgcagcag gcctaagagt gctctgttg agatcaaga aagaatcaa aagaccaag  
480  
atgaaagaa ggcctaagag gtggaattg ctctaaaga acagaaggtc aaggttaatt  
540  
tcccaagc tgcgtctga tccaagggtc ctaagtggtt agtggtggtt ggcacaagct  
602  
ga

&lt;210&gt; 321

&lt;211&gt; 108

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..108:Ceres Seq. ID 1709971

&lt;400&gt;321

Met Tyr Arg Lys Gln Xaa Lys Lys Asp Ala Ala Gln Glu Ala Val Lys  
1 5 10 15  
Arg Arg Arg Arg Ala Thr Lys Lys Pro Tyr Ser Arg Ser Ile Val Gly  
20 25 30  
Ala Thr Leu Glu Val Ile Gln Lys Lys Arg Ala Glu Lys Pro Glu Val  
35 40 45

Arg Asp Ala Ala Arg Glu Ala Ala Leu Arg Glu Ile Lys Glu Arg Ile  
50 55 60  
Lys Lys Thr Lys Asp Glu Lys Lys Ala Lys Lys Val Glu Phe Ala Ser  
65 70 75 80  
Lys Gln Gln Lys Val Lys Ala Asn Phe Pro Lys Ala Ala Ala Ala Ser  
85 90 95  
Lys Gly Pro Lys Val Gly Gly Gly Lys Arg  
100 105

&lt;210&gt; 322

&lt;211&gt; 59

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..59:Ceres Seq. ID 1709972

&lt;400&gt;322

Met Gln Pro Gly Lys Leu Leu Cys Val Arg Ser Arg Lys Glu Ser Lys  
1 5 10 15  
Arg Pro Lys Met Lys Arg Arg Leu Arg Arg Trp Asn Leu Leu Ser  
20 25 30  
Asn Arg Arg Ser Arg Leu Ile Ser Pro Lys Leu Leu His Pro Arg  
35 40 45  
Val Leu Arg Trp Glu Val Val Ala Asn Ala  
50 55

&lt;210&gt; 323

&lt;211&gt; 726

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..726:Ceres Seq. ID 1711273

&lt;400&gt;323

atctcttgt ctctcttca ttcttgta ttatccaga acgaagaaa acctagaaaa  
60  
cagttgaaga aagaaatca caagagaagc catggccga attgaccca ttactcagga  
120  
ttgggaacca gtgtgatcc gaagagagc tcttaagct gcagctaac ggcagcgaa  
180  
gactgcaac gccgtcgtc gaagcgccg cgtattgag accttgaa aattcaatgc  
240  
tggatgaaac aagctgcat caagcgccac ctcttgac acaagaagc tagatga  
300  
tactggaac ttatctcatg atcgtgccc cactgaattg aagaagcca tcatgcaagc  
360  
tagaggggaa aaaaagctga ctcatccca acttgccat ctgacatg agagccaca  
420  
agtgtacca gaatacagt cggggaagc aattccaat caacagatc ttccaagct  
480  
ggagaggcca ctgttgcta aactcgttg aagaagtag aagtttaga caaagctctt  
540  
aaagtaaca aaaagctgat cgcagttct ctccagcca catgcttac catacctaa  
600  
aaactatc tatgtatgt ttggttaat ggcgtagtag ttgttgca ggaatcttc  
660  
atgatgaag aaaaacaag ctgttgga cttttgca ttataataa tctctctct  
720  
ttcttt  
726

&lt;210&gt; 324

&lt;211&gt; 142

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown



<223> LOCATION: 1..142:Ceres Seq. ID 1711274

<400>324

Met Ala Gly Ile Gly Pro Ile Thr Gln Asp Trp Glu Pro Val Val Ile  
1 5 10 15  
Arg Lys Arg Ala Pro Asn Ala Ala Lys Arg Asp Glu Lys Thr Val  
20 25 30  
Asn Ala Ala Arg Arg Ser Gly Ala Asp Ile Glu Thr Val Arg Lys Phe  
35 40 45  
Asn Ala Gly Ser Asn Lys Ala Ala Ser Ser Gly Thr Ser Leu Asn Thr  
50 55 60  
Lys Lys Leu Asp Asp Thr Glu Asn Leu Ser His Asp Arg Val Pro  
65 70 75 80  
Thr Glu Leu Lys Lys Ala Ile Met Gln Ala Arg Gly Glu Xaa Lys Leu  
85 90 95  
Thr Gln Ser Gln Leu Ala His Leu Ile Asn Glu Lys Pro Gln Val Ile  
100 105 110  
Gln Glu Tyr Glu Ser Gly Lys Ala Ile Pro Asn Gln Ile Leu Ser  
115 120 125  
Lys Leu Glu Arg Ala Leu Gly Ala Lys Leu Arg Gly Lys Lys  
130 135 140

<210> 325

<211> 686

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..686:Ceres Seq. ID 1715423

<400>325

aacataaaaa caaagtgtt tatttgccg ttaagagaga tgataataga tgaagtgtat  
60  
gacatagaga gtacacagta tcaactctt tgccttgata aagccaagga gctcttgtct  
120  
ctctcaaac taaccacggg attgttgcca tgaagagata tgaaggagt tgggtacaac  
180  
aaucaaaa gggttggtg gatgggggtg agaagcaaga ttgagcatc attcgtgag  
240  
ataggtcgca gaggtttata tgaacgtgta gataactgac ttgtttgagg accgccgtat  
300  
gagagacct accggagta aagcaagaa gctcatgac ttgttctct tgaatgatat  
360  
cttctcaaa gagaagatc ctgagaagat cactttgct aatccacgg cctgtcacg  
420  
aacatttaa gtttcagcat tcaatgaga aggttgataa ataaagaa aagaggagta  
480  
accgacttc ctgatgaaa gcaaccaata aagaacagaa gtgttccat actcaataag  
540  
agatagttta attaatatc aaggacacgc cttcttaca atgaattgct tatggatcat  
600  
cacagtaaat aaacggttc tgaattgat gctctgatgt ataacagiga gtttcaatc  
660  
aacatatagg aatcagatca ttggtt

<210> 326

<211> 87

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..87:Ceres Seq. ID 1715424

<400>326

Met Ala Leu Arg Glu Met Ile Ile Asp Glu Ser Asp Asp Ile Glu Ser  
1 5 10 15  
Tyr Ser Asp Gln Ser Leu Cys Leu Asp Lys Ala Lys Glu Leu Leu Ala

20 25 30

Leu Ile Lys Leu Pro Thr Gly Leu Leu Pro Leu Lys Asp Met Thr Glu  
35 40 45  
Val Gly Tyr Asn Lys Thr Lys Thr Phe Val Trp Met Arg Leu Arg Ser  
50 55 60  
Lys Ile Glu His Thr Phe Arg Glu Ile Gly Arg Arg Val Leu Tyr Asp  
65 70 75 80  
Thr Xaa Asp Asn Cys Val Cys  
85

<210> 327

<211> 82

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..82:Ceres Seq. ID 1715425

<400>327

Met Ile Ile Asp Glu Ser Asp Asp Ile Glu Ser Tyr Ser Asp Gln Ser  
1 5 10 15  
Leu Cys Leu Asp Lys Ala Lys Glu Leu Leu Ala Leu Ile Lys Leu Pro  
20 25 30  
Thr Gly Leu Leu Pro Leu Lys Asp Met Thr Glu Val Gly Tyr Asn Lys  
35 40 45  
Thr Lys Gly Phe Val Trp Met Arg Leu Arg Ser Lys Ile Glu His Thr  
50 55 60  
Phe Arg Glu Ile Gly Arg Arg Val Leu Tyr Asp Thr Xaa Asp Asn Cys  
65 70 75 80  
Val Cys

<210> 328

<211> 65

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..65:Ceres Seq. ID 1715426

<400>328

Met Thr Arg Xaa Ile Thr Ala Phe Val Glu Asp Arg Arg Met Arg Arg  
1 5 10 15  
Leu Thr Gly Val Lys Ser Lys Glu Leu Met Ile Trp Val Pro Val Asn  
20 25 30  
Asp Ile Phe Ile Lys Glu Lys Asp Pro Glu Lys Ile Thr Phe Ala Asn  
35 40 45  
Thr Thr Gly Leu Ser Arg Thr Phe Lys Val Ser Ala Phe Gln Cys Glu  
50 55 60  
Gly  
65

<210> 329

<211> 829

<212> DNA

<213> Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..829:Ceres Seq. ID 1715962

&lt;400&gt;329

aaaaagccg tcgagaatc tccctcagg gttctttga tcaaacaga gagcgtcag 60  
 gagagcaaa acaagatcg gagaatga agcagttct tcttccgaa acgatggaca 120  
 tcccagacg tgtaccatc aagttcacg ctaagtgat cgaagtcaa ggaactgcg 180  
 ggaagcttg tccgatttc agcatctca acccgattt ccagctgatc aggatccag 240  
 agactgaaa gaagagctt agatcgatt cgtggttgg aaacgcaa accagcgct 300  
 ccacgaacg cgtcttagc cagctgata actgtatc cgtgttacc agagtttcc 360  
 gttacaagt gaggttcg tagccattt tcccatcaa cgcctccatc ggcgtgacg 420  
 gaaagtctat cgaatccgt aactccctg gggagaaga ggtgaggaag gtagagatg 480  
 tggatgtgt aacCattgt cgaYcTgaga aggtgaaga tgaatgttt cttagcgtta 540  
 acgacatcaa gctgtttca agtcatcgc cttgatcaa ccagaatgt cactggaaga 600  
 ggaagatag caggaggttt ctTgatgga tctatgttag cgaagaaagc aagatcgtag 660  
 aggaagaag aatgcccata tcatcgttt agtctctatt tctgtcttt tgaatgttag 720  
 ttctgttta tggaccaaat ccaccgtgt tgcaaaactc tgcataccc ttgttccctt 780  
 tttgtgtcc agattttatt taactatga caagttttgg agacagct 829

&lt;210&gt; 330

&lt;211&gt; 194

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..194:Ceres Seq. ID 1715963

&lt;400&gt;330

Met Lys Thr Ile Leu Ser Ser Glu Thr Met Asp Ile Pro Asp Ser Val 1  
 1 Thr Ile Lys Val His Ala Lys Val Ile Glu Val Glu Gly Pro Arg Gly 15  
 20 Lys Leu Val Arg Asp Phe Lys His Leu Asn Leu Asp Phe Gln Leu Ile 25  
 30 Lys Asp Pro Glu Thr Gly Lys Lys Leu Lys Ile Asp Ser Trp Phe 35  
 40 50 Gly Thr Arg Lys Thr Ser Ala Ser Ile Arg Thr Ala Leu Ser His Val 55  
 60 65 Asp Asn Leu Ile Ser Gly Val Thr Arg Gly Phe Arg Tyr Lys Met Arg 75  
 80 85 Phe Val Tyr Ala His Phe Pro Ile Asn Ala Ser Ile Gly Gly Asp Gly 90  
 95 Lys Ser Ile Glu Ile Arg Asn Phe Leu Gly Glu Lys Lys Val Arg Lys 100  
 105 Val Glu Met Leu Asp Gly Val Thr Ile Val Arg Xaa Glu Lys Val Lys 110  
 115 120 125 Asp Glu Ile Val Leu Asp Gly Asn Asp Ile Glu Leu Val Ser Arg Ser 130  
 135 140 Cys Ala Leu Ile Asn Gln Lys Cys His Val Lys Lys Asp Ile Arg 145  
 150 155 Lys Phe Leu Asp Gly Ile Tyr Val Ser Glu Lys Ser Lys Ile Val Glu 160  
 165 170 175 Glu Glu 180

&lt;210&gt; 331

&lt;211&gt; 185

159

SUBSTITUTE SHEET (RULE 26)

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..185:Ceres Seq. ID 1715964

&lt;400&gt;331

Met Asp Ile Pro Asp Ser Val Thr Ile Lys Val His Ala Lys Val Ile 1  
 5 10 15 Glu Val Glu Gly Pro Arg Gly Lys Leu Val Arg Asp Phe Lys His Leu 20  
 25 30 Asn Leu Asp Phe Gln Leu Ile Lys Asp Pro Glu Thr Gly Lys Lys Lys 35  
 40 45 Leu Lys Ile Asp Ser Trp Phe Gly Thr Arg Lys Thr Ser Ala Ser Ile 50  
 55 60 Arg Thr Ala Leu Ser His Val Asp Asn Leu Ile Ser Gly Val Thr Arg 65  
 70 75 Gly Phe Arg Tyr Lys Met Arg Phe Val Tyr Ala His Phe Pro Ile Asn 80  
 85 90 Ala Ser Ile Gly Gly Asp Gly Lys Ser Ile Glu Ile Arg Asn Phe Leu 95  
 100 105 Gly Glu Lys Lys Val Arg Lys Val Glu Met Leu Asp Gly Val Thr Ile 110  
 115 120 Val Arg Xaa Glu Lys Val Lys Asp Glu Ile Val Leu Asp Gly Asn Asp 125  
 130 135 Ile Glu Leu Val Ser Arg Ser Cys Ala Leu Ile Asn Gln Lys Cys His 140  
 145 150 Val Lys Lys Lys Asp Ile Arg Lys Phe Leu Asp Gly Ile Tyr Val Ser 155  
 160 165 Glu Lys Ser Lys Ile Val Glu Glu Glu 170  
 175 180

&lt;210&gt; 332

&lt;211&gt; 100

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..100:Ceres Seq. ID 1715965

&lt;400&gt;332

Met Arg Phe Val Tyr Ala His Phe Pro Ile Asn Ala Ser Ile Gly Gly 1  
 5 10 15 Asp Gly Lys Ser Ile Glu Ile Arg Asn Phe Leu Gly Glu Lys Lys Val 20  
 25 30 Arg Lys Val Glu Met Leu Asp Gly Val Thr Ile Val Arg Xaa Glu Lys 35  
 40 45 Val Lys Asp Glu Ile Val Leu Asp Gly Asn Asp Ile Glu Leu Val Ser 50  
 55 60 Arg Ser Cys Ala Leu Ile Asn Gln Lys Cys His Val Lys Lys Lys Asp 65  
 70 75 Ile Arg Lys Phe Leu Asp Gly Ile Tyr Val Ser Glu Lys Ser Lys Ile 80  
 85 90 Val Glu Glu Glu 100

&lt;210&gt; 333

160

SUBSTITUTE SHEET (RULE 26)

<211> 675  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..675:Ceres Seq. ID 1808584

<400>333  
aactgtcac ctgcagagaa agaagaagc cctagattt gtcaaaagc gggtgcagaa 60  
caaaaacca ttgcccgaat tgaccgata actcaggatt ggggcccgtt ggtgatccgt 120  
aagaaccgc tcaacgcgc tgcagagcg gcagagaata ctgtcaaacgc ggtgcctga 180  
tcgggcccq atagcagac cgtcagaata tcaatgctg gaaccaacaa gggcgatca 240  
agcgcaact ctgtgaacac aanaatgctt gttgatgac ctgagaacct tactatgaa 300  
cgtgtccaa ttgtcctaa gaaagccatt atgcagcca ggaagacaa gaagctaac 360  
cagtcacaac ttgtcaaat catcaatgag aagccaag tgattcaaga gtatgagtt 420  
ggcaagcta taccacaaca gcaatcctt tctaagctgg agagagcgt tggagctag 480  
cttcgggaa agaagtgagc caagttctac tgaatgagca agtaacaga atcaatgctt 540  
tcgtcaatg ccgtaaactt gccagaaga atatttctg attgaagaa agcaaacgc 600  
ttgaatgtt tgttcgttg atgaatctc tatctcataa actcatatca atataatac 660  
ttgggtcttt tcatc 675

<210> 334  
<211> 165  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..165:Ceres Seq. ID 1808585

<400>334  
Asn Leu Ser Pro Ala Glu Lys Glu Arg Ser Pro Arg Phe Cys Gln Lys 5  
1 Ala Val Ala Glu Gln Lys Thr Met Ala Gly Ile Gly Pro Ile Thr Gln 15  
20 25 30  
Asp Trp Glu Pro Val Val Ile Arg Lys Lys Pro Ala Asn Ala Ala Ala 35  
35 40 45  
Lys Arg Asp Glu Lys Thr Val Asn Ala Ala Arg Arg Ser Gly Ala Asp 50  
50 55 60  
Ile Glu Thr Val Arg Lys Phe Asn Ala Gly Thr Asn Lys Ala Ala Ser 65  
65 70 75  
Ser Gly Thr Ser Leu Asn Thr Lys Met Leu Asp Asp Thr Glu Asn 80  
85 90 95  
Leu Thr His Glu Arg Val Pro Thr Glu Leu Lys Lys Ala Ile Met Gln 100  
105 110  
Ala Arg Thr Asp Lys Lys Leu Thr Gln Ser Gln Leu Ala Gln Ile Ile 115  
120 125  
Asn Glu Lys Pro Gln Val Ile Gln Glu Tyr Glu Ser Gly Lys Ala Ile 130  
135 140  
Pro Asn Gln Gln Ile Leu Ser Lys Leu Glu Arg Ala Leu Gly Ala Lys 145  
150 155 160  
Leu Arg Gly Lys Lys 165

<210> 335  
<211> 142  
<212> PRT  
<213> Arabidopsis thaliana

161

SUBSTITUTE SHEET (RULE 26)

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..142:Ceres Seq. ID 1808586

<400>335  
Met Ala Gly Ile Gly Pro Ile Thr Gln Asp Trp Glu Pro Val Val Ile 15  
1 Arg Lys Lys Pro Ala Asn Ala Ala Lys Arg Asp Glu Lys Thr Val 30  
20 25 30  
Asn Ala Ala Arg Arg Ser Gly Ala Asp Ile Glu Thr Val Arg Lys Phe 35  
35 40 45  
Asn Ala Gly Thr Asn Lys Ala Ala Ser Ser Gly Thr Ser Leu Asn Thr 50  
50 55 60  
Lys Met Leu Asp Asp Thr Glu Asn Leu Thr His Glu Arg Val Pro 65  
65 70 75  
Thr Glu Leu Lys Lys Ala Ile Met Gln Ala Arg Thr Asp Lys Lys Leu 80  
85 90 95  
Thr Gln Ser Gln Leu Ala Gln Ile Ile Asn Glu Lys Pro Gln Val Ile 100  
105 110  
Gln Glu Tyr Glu Ser Gly Lys Ala Ile Pro Asn Gln Gln Ile Leu Ser 115  
120 125  
Lys Leu Glu Arg Ala Leu Gly Ala Lys Leu Arg Gly Lys Lys 130  
135 140

<210> 336  
<211> 630  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..630:Ceres Seq. ID 1808591

<400>336  
aaactcaaa caatttcoat acctcatatt ctcttcacg gcaaccacca agcaactcct 60  
ctcttcctc ctcatcatcg ctatcatcac cattacatct tccacttcac ttccatttct 120  
cacgcagag caaaaccaaa tgcacaacaa aatcatagac gcaatggtct caagtggctc 180  
tttcgaagat tggagcggag cgtctctcaa caacaacgac gaattaaacg gtccagtctt 240  
aactcaact ctcttctcc ccaaaacatc cgtggagga atcaacgcca cgtccaccat 300  
agttgtctct taccatattg ttccacaatg gcttgacttc tccgctataa gtctcatgat 360  
gcctttctc cgcataccta cactctctc tggacactct atagctgta ccaacaatic 420  
agcttctggt ttactcttg atggtgttct catctctgag ccagatttat tegtctctcc 480  
tactatagt atccatcgaa tggcttttcc attaaacttc tccgtttacg gggggcgga 540  
tatatagta ttactcttg atcgcttct ccactttgt tctttctact ttactaata 600  
cacttcaaa taattacaa cgaacttgac 630

<210> 337  
<211> 181  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..181:Ceres Seq. ID 1808592

<400>337  
Thr Leu Gln Gln Phe Ser Tyr Leu Ile Phe Ser Ser Met Ala Thr Thr

162

SUBSTITUTE SHEET (RULE 26)

1 Iys His Leu Leu Leu Phe Leu Leu Ile Ile Ala Ile Ile Thr Ile Thr 15  
20 23 30  
Ser Ser Thr Leu Leu Pro Phe Leu Thr Thr Glu Gln Asn Gln Ile Ala 45  
35 40 45  
Thr Lys Ile Ile Asp Ala Met Val Ser Ser Gly Ser Phe Glu Asp Trp 60  
50 55 60  
Ser Gly Ala Phe Leu Asn Asn Asp Glu Leu Asn Gly Pro Val Leu 75  
65 70 75 80  
Thr Ser Thr Leu Phe Leu Pro Lys Thr Ser Val Glu Gly Ile Asn Ala 95  
85 90 95  
Thr Ser Pro Leu Val Ala Ser Tyr Tyr His Ile Val Pro Gln Trp Leu Asp 110  
Phe Ser Val Ile Ser Leu Met Met Pro Phe Ser Arg Ile Pro Thr Leu 125  
115 120 125  
Leu Ser Gly His Ser Ile Val Val Thr Asn Asn Ser Ala Ser Gly Phe 140  
130 135 140  
Thr Leu Asp Gly Val Leu Ile Ser Glu Pro Asp Leu Phe Val Ser Pro 155  
145 150 155 160  
Thr Ile Val Ile His Arg Met Ala Phe Pro Phe Asn Phe Ser Arg Tyr 175  
165 170 175  
Gly Gly Gly Asp Ile 180

<210> 338  
<211> 169  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..169:Ceres Seq. ID 1808593

<400>338  
Met Ala Thr Thr Lys His Leu Leu Leu Phe Leu Leu Ile Ile Ala Ile 15  
1 5 10 15  
Ile Thr Ile Thr Ser Ser Thr Ser Leu Pro Phe Leu Thr Thr Glu Gln 20  
20 25 30  
Asn Gln Ile Ala Thr Lys Ile Ile Asp Ala Met Val Ser Ser Gly Ser 45  
35 40 45  
Phe Glu Asp Trp Ser Gly Ala Phe Leu Asn Asn Asn Asp Glu Leu Asn 50  
55 60  
Gly Pro Val Leu Thr Ser Thr Leu Phe Leu Pro Lys Thr Ser Val Glu 75  
65 70 75  
Gly Ile Asn Ala Thr Ser Pro Leu Val Ala Ser Tyr Tyr His Ile Val Pro 95  
85 90 95  
Gln Trp Leu Asp Phe Ser Val Ile Ser Leu Met Met Pro Phe Ser Arg 110  
100 105 110  
Ile Pro Thr Leu Leu Ser Gly His Ser Ile Val Val Thr Asn Asn Ser 125  
115 120 125  
Ala Ser Gly Phe Thr Leu Asp Gly Val Leu Ile Ser Glu Pro Asp Leu 140  
130 135 140  
Phe Val Ser Pro Thr Ile Val Ile His Arg Met Ala Phe Pro Phe Asn 155  
145 150 155 160  
Phe Ser Arg Tyr Gly Gly Asp Ile 165  
165 165

<210> 339  
<211> 127  
<212> PRT

163

SUBSTITUTE SHEET (RULE 26)

<213> Arabidopsis thaliana  
<220>  
<223> any n or Xaa = unknown  
<223> LOCATION: 1..127:Ceres Seq. ID 1808594  
<400>339  
Met Val Ser Ser Gly Ser Phe Glu Asp Trp Ser Gly Ala Phe Leu Asn 15  
1 5 10 15  
Asn Asn Asp Glu Leu Asn Gly Pro Val Leu Thr Ser Thr Leu Phe Leu 30  
20 25 30  
Pro Lys Thr Ser Val Glu Gly Ile Asn Ala Thr Ser Pro Leu Val Ala 45  
35 40 45  
Ser Tyr His Ile Val Pro Gln Trp Leu Asp Phe Ser Val Ile Ser Leu 60  
50 55 60  
Met Met Pro Phe Ser Arg Ile Pro Thr Leu Leu Ser Gly His Ser Ile 75  
65 70 75 80  
Val Val Thr Asn Asn Ser Ala Ser Gly Phe Thr Leu Asp Gly Val Leu 95  
Ile Ser Glu Pro Asp Leu Phe Val Ser Pro Thr Thr Ile Val Ile His Arg 110  
100 105 110  
Met Ala Phe Pro Phe Asn Phe Ser Arg Tyr Gly Gly Asp Ile 125  
115 120 125

<210> 340  
<211> 717  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..717:Ceres Seq. ID 1920563

<400>340  
ctgtgtata ctaatgacca agagagagact gtaaaaggag ttgtctcaaa ttcttgtgt 60  
gatttggcac ccgaagtgga tgtaagtgt actggtctctg taggcaaaaga aatgcttatg 120  
cctaagatc caaagccac cgttaattatg ctgcccacag ggaacagat tgcctcttc 180  
aggctttct tatggaagt gtcttttgag aaacatgatg actacaagt caatggctta 240  
gcttggtgt tcttggtgt accaaccact agctcattgc tctaccaaga ggaattgat 300  
aagatgaag caaaggccc ccgaaacttc aggttggtt acgcgataag cagagacaa 360  
gcgaacgata aaggagaaa aatgtatatc cagactcgga tggcacagta cgcagctgaa 420  
ttatggagt tgtgaagaa agacaacact ttgtttaca tgttggaact caagggaatg 480  
gagaaggaaa ttgatgacat tatggtctca ttggtgcaa atgaaggat tgaatggttt 540  
gattacaaga agcagttgaa gaaggcagag caatggaacg ttgaagtcta ctgatcaaa 600  
agcctttgac attctgtag caaagtatag ctgaacaaa ctgtaatttt cgcctccgaa 660  
ttctgtatt tgaagataa gttttttaga tatgtttatc taaaaaaga gttcttt 717

<210> 341  
<211> 197  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..197:Ceres Seq. ID 1920564

<400>341  
Leu Val Tyr Thr Asn Asp Gln Gly Glu Thr Val Lys Lys Gly Val Cys Ser

164

SUBSTITUTE SHEET (RULE 26)

1 Asn Phe Leu Cys Asp Leu Ala Pro Gly Ser Asp Val Lys Leu Thr Gly 15  
 20 25 30  
 Pro Val Gly Lys Glu Met Leu Met Pro Lys Asp Pro Asn Ala Thr Val  
 35 40 45  
 Ile Met Leu Ala Thr Gly Thr Gly Ile Ala Pro Phe Arg Ser Phe Leu  
 50 55 60  
 Trp Lys Met Phe Phe Glu Lys His Asp Asp Tyr Lys Phe Asn Gly Leu  
 65 70 75 80  
 Ala Trp Leu Phe Leu Gly Val Pro Thr Thr Ser Ser Leu Leu Tyr Gln  
 85 90 95  
 Glu Glu Phe Asp Lys Met Lys Ala Lys Ala Pro Glu Asn Phe Arg Val  
 100 110 115  
 Asp Tyr Ala Ile Ser Arg Glu Gln Ala Asn Asp Lys Gly Glu Lys Met  
 120 125  
 Tyr Ile Gln Thr Arg Met Ala Gln Tyr Ala Ala Glu Leu Trp Glu Leu  
 130 135 140  
 Leu Lys Lys Asp Asn Thr Phe Val Tyr Met Cys Gly Leu Lys Gly Met  
 145 150 155 160  
 Glu Lys Gly Ile Asp Asp Ile Met Val Ser Leu Ala Ala Asn Asp Gly  
 165 170 175  
 Ile Asp Trp Phe Asp Tyr Lys Lys Gln Leu Lys Lys Ala Glu Gln Trp  
 180 185 190  
 Asn Val Glu Val Tyr 195

&lt;210&gt; 342

&lt;211&gt; 160

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..160:Ceres Seq. ID 1920565

<400>342  
 Met Leu Met Pro Lys Asp Pro Asn Ala Thr Val Ile Met Leu Ala Thr  
 1 5 10 15  
 Gly Thr Gly Ile Ala Pro Phe Arg Ser Phe Leu Trp Lys Met Phe Phe  
 20 25 30  
 Glu Lys His Asp Asp Tyr Lys Phe Asn Gly Leu Ala Trp Leu Phe Leu  
 35 40 45  
 Gly Val Pro Thr Thr Ser Ser Leu Leu Tyr Gln Glu Glu Phe Asp Lys  
 50 55 60  
 Met Lys Ala Lys Ala Pro Glu Asn Phe Arg Val Asp Tyr Ala Ile Ser  
 65 70 75 80  
 Arg Glu Gln Ala Asn Asp Lys Gly Glu Lys Met Tyr Ile Gln Thr Arg  
 85 90 95  
 Met Ala Gln Tyr Ala Ala Glu Leu Trp Glu Leu Lys Lys Asp Asn  
 100 105 110  
 Thr Phe Val Tyr Met Cys Gly Leu Lys Gly Met Glu Lys Gly Ile Asp  
 115 120 125  
 Asp Ile Met Val Ser Leu Ala Ala Asn Asp Gly Ile Asp Trp Phe Asp  
 130 135 140  
 Tyr Lys Lys Gln Leu Lys Lys Ala Glu Gln Trp Asn Val Glu Val Tyr  
 145 150 155 160

&lt;210&gt; 343

&lt;211&gt; 158

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..158:Ceres Seq. ID 1920566

<400>343  
 Met Pro Lys Asp Pro Asn Ala Thr Val Ile Met Leu Ala Thr Gly Thr  
 1 5 10 15  
 Gly Ile Ala Pro Phe Arg Ser Phe Leu Trp Lys Met Phe Phe Glu Lys  
 20 25 30  
 His Asp Asp Tyr Lys Phe Asn Gly Leu Ala Trp Leu Phe Leu Gly Val  
 35 40 45  
 Pro Thr Thr Ser Ser Leu Leu Tyr Gln Glu Glu Phe Asp Lys Met Lys  
 50 55 60  
 Ala Lys Ala Pro Glu Asn Phe Arg Val Asp Tyr Ala Ile Ser Arg Glu  
 65 70 75 80  
 Gln Ala Asn Asp Lys Gly Glu Lys Met Tyr Ile Gln Thr Arg Met Ala  
 85 90 95  
 Gln Tyr Ala Ala Glu Leu Trp Glu Leu Lys Lys Asp Asn Thr Phe  
 100 105 110  
 Val Tyr Met Cys Gly Leu Lys Gly Met Glu Lys Gly Ile Asp Asp Ile  
 115 120 125  
 Met Val Ser Leu Ala Ala Asn Asp Gly Ile Asp Trp Phe Asp Tyr Lys  
 130 135 140  
 Lys Gln Leu Lys Lys Ala Glu Gln Trp Asn Val Glu Val Tyr  
 145 150 155

&lt;210&gt; 344

&lt;211&gt; 2192

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..2192:Ceres Seq. ID 1974419

<400>344  
 attcaagtc attctatctc gtcaacaata caaaccaacc ttctcaattc ctctcttttc  
 attccatttt tcaatagcta atggcttcta gtccaatgag caactcata tggacagcta  
 gagaggacaa gcaattcgaa atggctgttg cgaatttoga caaggacact cctgaccgtt  
 ggcacaaat tgcagggca gttgtgtgga atcaactga agaagtaaa cgaactatg  
 attgtccct tagggatgtg aatgacattg agtcaggcgc ctatccaaa cctagattat  
 gtccggggc ggcgatgct gtgcgttca ctccaatgc gatgtcacag aaccgctt  
 acgtgactc gsgatttcgc ttctcttct tctctcttc acctccgca gaagaatc  
 caaccgatac taagtcttct tgaattctta aatcgagac taaacctgac tccgtatgac  
 caaagatc tggttctgct cctgattcat tgaagaggg tgetaaagct tctcgtgaa  
 tcaatagctc tctctatcc aaacaggtgt ttgatcta gcgaagcag gagagact  
 gattagctga attagcgtt ggaagaaac ataatgaac tattcaagt agcaaggca  
 tgaagaca ggcgaattg gcagagatc agagaattt agtcagcaa caggcccaag  
 cgaagcgca aaattctaga tatgagatg agttggccag gaagagatg cagacagata  
 tgaagctca gagacgcat aatgctgaat tggttcgt gcaagagca tctctatc  
 ggaagagaa agcaagaatt gccacggag aacagattca agcacagcag cgcgagactg  
 agaaagagag agcggaact gagcgagaa caattcgct gaaagctat gctgaggtg  
 agggccgagc tcatgaagct aaactcagc aggcagaa tagaagatg ctctagata  
 agataaatgg tgaagggag aaatggcttg cagcaatca cagacttcc agtcacattg  
 agagagatt aggcctta ttaactgac gaagtaatt gattagct gtggagag  
 ttacggcatt agccgtggg gttaacaa ctcgaaggg tctagggtt acctgggtt  
 atatcaatag aattctggg cagccatcac tgatcagaga atcttccatg ggcgattcc

catggcagc ctacgtgtct cagtttaaga aaaaacttag cacagctgca ggggcagcag  
 catctcaga agagaaaaag cctcttgaaa atgtaattct coactgttct ttgaagacga  
 gaattgagc ttctgcaaga gccacagcaa ataccaagtc acataaagca ccattccgaa  
 acatgatgt ttatggcct ccaggtaccg gaanaactat ggtgcgagc gagattgtct  
 ggaagtcgg ttctgatat gctatgaga caggaggaga tgttctcct ctgtgtcac  
 aggtgttac aagaatccat gaatatattg attggctaa gaatacaaac aaagggttac  
 tgccttcat cgatgaagct gatgtcttc tatgcgaacg taacagcaact tacatggtg  
 aggtcagcg cagcgtctg aacgcttgc tcttgcgac cggtgatcaa tgcgggaca  
 1740 1800 1860 1920 1980 2040 2100 2160 2192  
 tagctctgt cfcggtaca aacagactg gagatctga cagtgacgc actgacaga  
 tgcagagt tatcaggttt cctctccgt gtagaaga acgttcaag ctcctcaagc  
 tctatctcaa ctgtacctta atggtgacg acaagaagc tgagaagac tcaacctta  
 aatgagcaa ctgttctcaag aagaagatc cacagaagt aaccttgaa ggagacctaa  
 ccgaccaagt gattaagaa gctgcaaaa agacagaagc cttttcggc cgtgaacgc  
 ctacgttgt cgcggttgt cagctgcgg tatatggacg acaggattgt gcttggatt  
 cccaacttt tgaagaatt gtgagtata agatcgaaga acatcaccag aagaatcac  
 ttgcgactga aggtggccaa tctttccgt ag

&lt;210&gt; 345

&lt;211&gt; 703

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..703:Ceres Seq. ID 1974420

&lt;400&gt;345

Met Ala Ser Ser Ser Met Ser Thr Ser Thr Ala Arg Glu Asp  
 1 5 10 15  
 Lys Gln Phe Glu Met Ala Leu Ala Lys Phe Asp Lys Asp Thr Pro Asp  
 20 25 30  
 Arg Trp Gln Lys Ile Ala Arg Ala Val Gly Lys Ser Thr Thr Glu Glu  
 35 40 45  
 Val Lys Arg His Tyr Glu Leu Leu Arg Asp Val Asn Asp Ile Glu  
 50 55 60  
 Ser Gly Arg Tyr Pro Gln Pro Arg Leu Cys Ser Ala Ala Ile Ala  
 65 70 75 80  
 Ala Ala Phe Thr Ser Met Ser Met Ser Gln Asn Arg Ala Tyr Ala Asp  
 85 90 95  
 Ser Arg Phe Arg Phe Pro Phe Ser Ser Ser Pro Pro Ala Glu Glu  
 100 105 110  
 Ser Pro Thr Asp His Lys Ser Ser Ser Asn Ser Lys Ser Glu Thr Lys  
 115 120 125  
 Pro Asp Ser Asp Glu Pro Lys Gly Ser Gly Phe Asp Pro Glu Ser Leu  
 130 135 140  
 Glu Arg Gly Ala Lys Ala Leu Arg Glu Ile Asn Ser Ser Pro His Ser  
 145 150 155 160  
 Lys Gln Val Phe Asp Leu Met Arg Lys Gln Glu Lys Thr Arg Leu Ala  
 165 170 175  
 Glu Leu Ala Ala Glu Lys Glu His Asn Glu Ala Ile Gln Ala Ser Lys  
 180 185 190  
 Asp Ile Glu Arg Gln Arg Lys Leu Ala Glu Asp Gln Arg Asn Leu Val  
 195 200 205  
 Gln Gln Gln Ala Gln Ala Lys Ala Gln Asn Leu Arg Tyr Glu Asp Glu  
 210 215  
 Leu Ala Arg Lys Arg Met Gln Thr Asp Asn Glu Ala Gln Arg Arg His  
 225 230 235 240  
 Asn Ala Glu Leu Val Ser Met Gln Glu Ala Ser Ser Ile Arg Lys Glu  
 245 250 255  
 Lys Ala Arg Ile Ala Thr Glu Glu Gln Ile Gln Ala Gln Gln Arg Glu

Thr Glu Lys Glu Arg Ala Glu Leu Glu Arg Glu Thr Ile Arg Val Lys  
 260 265 270  
 Ala Met Ala Glu Ala Glu Gly Arg Ala His Glu Ala Lys Leu Thr Glu  
 275 280 285  
 Glu Gln Asn Arg Arg Met Leu Leu Asp Lys Ile Asn Gly Glu Arg Glu  
 290 295 300  
 Lys Trp Leu Ala Ala Ile Asn Thr Thr Phe Ser His Ile Glu Gly Gly  
 305 310 315 320  
 Val Arg Thr Leu Leu Thr Thr Asp Arg Ser Lys Leu Ile Met Thr Val Gly  
 320 325 330 335  
 Gly Val Thr Ala Leu Ala Ala Gly Val Tyr Thr Thr Arg Glu Gly Ala  
 335 340 345 350  
 Arg Val Thr Trp Gly Tyr Ile Asn Arg Ile Leu Gly Gln Pro Ser Leu  
 350 355 360 365  
 Ile Arg Glu Ser Ser Met Gly Arg Phe Pro Trp Ala Gly Ser Val Ser  
 365 370 375 380  
 Gln Phe Lys Asn Lys Leu Ser Thr Ala Ala Gly Ala Ala Ser Ala  
 380 385 390 395 400  
 Glu Gly Glu Lys Pro Leu Glu Asn Val Ile Leu His Arg Ser Leu Lys  
 400 405 410 415  
 Thr Arg Ile Glu Arg Leu Ala Arg Ala Thr Ala Asn Thr Lys Ser His  
 415 420 425 430  
 Lys Ala Pro Phe Arg Asn Met Met Phe Tyr Gly Pro Gly Thr Gly  
 430 435 440 445  
 Lys Thr Met Val Ala Arg Glu Ile Ala Arg Lys Ser Gly Leu Asp Tyr  
 445 450 455 460  
 Ala Met Met Thr Gly Gly Asp Val Ala Pro Leu Gly Ala Gln Ala Val  
 460 465 470 475 480  
 Thr Lys Ile His Glu Ile Phe Asp Trp Ala Lys Lys Ser Asn Lys Gly  
 480 485 490 495  
 Leu Leu Leu Phe Ile Asp Glu Ala Asp Ala Phe Leu Cys Glu Arg Asn  
 495 500 505 510  
 Ser Thr Tyr Met Ser Glu Ala Gln Arg Ser Ala Leu Asn Ala Leu Leu  
 510 515 520 525  
 Phe Arg Thr Gly Asp Gln Ser Arg Asp Ile Val Leu Val Leu Ala Thr  
 525 530 535 540  
 Asn Arg Pro Gly Asp Leu Asp Ser Ala Val Thr Asp Arg Ile Asp Glu  
 540 545 550 555  
 Val Ile Glu Phe Pro Leu Pro Gly Glu Glu Arg Phe Lys Leu Leu  
 555 560 565 570  
 Lys Leu Tyr Leu Asn Lys Tyr Leu Met Gly Asp Lys Lys Gly Glu  
 570 575 580 585 590  
 Lys Asp Ser Asn Leu Lys Trp Ser Asn Leu Phe Lys Lys Lys Ser  
 590 595 600 605  
 Gln Lys Ile Thr Ile Glu Gly Asp Leu Thr Asp Gln Val Ile Lys Glu  
 605 610 615 620  
 Ala Ala Lys Lys Thr Glu Gly Phe Ser Gly Arg Glu Ile Ala Lys Leu  
 620 625 630 635 640  
 Val Ala Gly Val Gln Ala Ala Val Tyr Gly Arg Gln Asp Cys Val Leu  
 640 645 650 655 660  
 Asp Ser Gln Leu Phe Glu Glu Ile Val Asp Tyr Lys Ile Glu Glu His  
 660 665 670 675  
 His Gln Arg Ile Arg Leu Ala Thr Glu Gly Gln Ser Phe Pro  
 675 680 685 690 695 700

&lt;210&gt; 346

&lt;211&gt; 698

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..698:Ceres Seq. ID 1974421

&lt;400&gt;346

Met Ser Ser Ser Trp Thr Ala Arg Glu Asp Lys Gln Phe Glu Met  
 1 5 10 15  
 Ala Leu Ala Lys Phe Asp Lys Asp Thr Pro Asp Arg Trp Gln Lys Ile  
 20 25 30  
 Ala Arg Ala Val Gly Gly Lys Ser Thr Glu Glu Val Lys Arg His Tyr  
 35 40 45  
 Glu Ileu Leu Leu Arg Asp Val Asn Asp Ile Glu Ser Gly Arg Tyr Pro  
 50 55 60  
 Gln Pro Arg Leu Cys Ser Ala Ala Ile Ala Ala Phe Thr Ser  
 65 70 75 80  
 Met Ser Met Ser Gln Asn Arg Ala Tyr Ala Asp Ser Arg Phe Arg Phe  
 85 90 95  
 Pro Phe Phe Ser Ser Pro Pro Ala Glu Glu Ser Pro Thr Asp His  
 100 105 110  
 Lys Ser Ser Ser Asn Ser Lys Ser Glu Thr Lys Pro Asp Ser Asp Glu  
 115 120 125  
 Pro Lys Gly Ser Gly Phe Asp Pro Glu Ser Leu Glu Arg Gly Ala Lys  
 130 135 140  
 Ala Leu Arg Glu Ile Asn Ser Ser Pro His Ser Lys Gln Val Phe Asp  
 145 150 155 160  
 Leu Met Arg Lys Gln Glu Lys Thr Arg Leu Ala Glu Leu Ala Ala Glu  
 165 170 175  
 Lys Glu His Asn Glu Ala Ile Gln Ala Ser Lys Asp Ile Glu Arg Gln  
 180 185 190  
 Arg Lys Leu Ala Glu Asp Gln Arg Asn Leu Val Gln Gln Gln Ala Gln  
 195 200 205  
 Ala Lys Ala Gln Asn Leu Arg Tyr Glu Asp Glu Leu Ala Arg Lys Arg  
 210 215 220  
 Met Gln Thr Asp Asn Glu Ala Gln Arg Arg His Asn Ala Glu Leu Val  
 225 230 235 240  
 Ser Met Gln Glu Ala Ser Ser Ile Arg Lys Glu Lys Ala Arg Ile Ala  
 245 250 255  
 Thr Glu Glu Gln Ile Gln Ala Gln Arg Glu Thr Glu Lys Glu Arg  
 260 265 270  
 Ala Glu Leu Glu Arg Glu Thr Ile Arg Val Lys Ala Met Ala Glu Ala  
 275 280 285  
 Glu Gly Arg Ala His Glu Ala Lys Leu Thr Glu Glu Gln Asn Arg Arg  
 290 295 300  
 Met Leu Leu Asp Lys Ile Asn Gly Glu Arg Glu Lys Trp Leu Ala Ala  
 305 310 315 320  
 Ile Asn Thr Thr Phe Ser His Ile Glu Gly Gly Val Arg Thr Leu Leu  
 325 330 335  
 Thr Asp Arg Ser Lys Leu Ile Met Thr Val Gly Gly Val Thr Ala Leu  
 340 345 350  
 Ala Ala Gly Val Tyr Thr Arg Glu Gly Ala Arg Val Thr Trp Gly  
 355 360 365  
 Tyr Ile Asn Arg Ile Leu Gly Gln Pro Ser Leu Ile Arg Glu Ser Ser  
 370 375 380  
 Met Gly Arg Phe Pro Trp Ala Gly Ser Val Ser Gln Phe Lys Asn Lys  
 385 390 395 400  
 Leu Ser Thr Ala Ala Gly Ala Ala Sor Ala Glu Gly Glu Lys Pro  
 405 410 415  
 Leu Glu Asn Val Ile Leu His Arg Ser Leu Lys Thr Arg Ile Glu Arg  
 420 425 430  
 Leu Ala Arg Ala Thr Ala Asn Thr Lys Ser His Lys Ala Pro Phe Arg

169

SUBSTITUTE SHEET (RULE 26)

435

440

445

Asn Met Met Phe Tyr Gly Pro Gly Thr Gly Lys Thr Met Val Ala  
 450 455 460  
 Arg Glu Ile Ala Arg Lys Ser Gly Leu Asp Tyr Ala Met Met Thr Gly  
 465 470 475 480  
 Gly Asp Val Ala Pro Leu Gly Ala Gln Ala Val Thr Lys Ile His Glu  
 485 490 495  
 Ile Phe Asp Trp Ala Lys Lys Ser Asn Lys Gly Leu Leu Phe Ile  
 500 505 510  
 Asp Glu Ala Asp Ala Phe Leu Cys Glu Arg Asn Ser Thr Tyr Met Ser  
 515 520 525  
 Glu Ala Gln Arg Ser Ala Leu Asn Ala Leu Leu Phe Arg Thr Gly Asp  
 530 535 540  
 Gln Ser Arg Asp Ile Val Leu Val Leu Ala Thr Asn Arg Pro Gly Asp  
 545 550 555 560  
 Leu Asp Ser Ala Val Thr Asp Arg Ile Asp Glu Val Ile Glu Phe Pro  
 565 570 575  
 Leu Pro Gly Glu Glu Arg Phe Lys Leu Leu Lys Leu Tyr Leu Asn  
 580 585 590  
 Lys Tyr Leu Met Gly Asp Asp Lys Lys Gly Glu Lys Asp Ser Asn Leu  
 595 600 605  
 Lys Trp Ser Asn Leu Phe Lys Lys Lys Lys Ser Gln Lys Ile Thr Ile  
 610 615 620  
 Glu Gly Asp Leu Thr Asp Gln Val Ile Lys Glu Ala Ala Lys Lys Thr  
 625 630 635 640  
 Glu Gly Phe Ser Gly Arg Glu Ile Ala Lys Leu Val Ala Gly Val Gln  
 645 650 655  
 Ala Ala Val Tyr Gly Arg Gln Asp Cys Val Leu Asp Ser Gln Leu Phe  
 660 665 670  
 Glu Glu Ile Val Asp Tyr Lys Ile Glu Glu His His Gln Arg Ile Arg  
 675 680 685  
 Leu Ala Thr Glu Gly Gln Ser Phe Pro  
 690 695

&lt;210&gt; 347

&lt;211&gt; 683

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..683:Ceres Seq. ID 1974422

&lt;400&gt;347

Met Ala Leu Ala Lys Phe Asp Lys Asp Thr Pro Asp Arg Trp Gln Lys  
 1 5 10 15  
 Ile Ala Arg Ala Val Gly Gly Lys Ser Thr Glu Glu Val Lys Arg His  
 20 25 30  
 Tyr Glu Leu Leu Leu Arg Asp Val Asn Asp Ile Glu Ser Gly Arg Tyr  
 35 40 45  
 Pro Gln Pro Arg Leu Cys Ser Ala Ala Ile Ala Ala Phe Thr  
 50 55 60  
 Ser Met Ser Met Ser Gln Asn Arg Ala Tyr Ala Asp Ser Arg Phe Arg  
 65 70 75  
 Phe Pro Phe Phe Ser Ser Pro Pro Ala Glu Glu Ser Pro Thr Asp  
 85 90 95  
 His Lys Ser Ser Ser Asn Ser Lys Ser Glu Thr Lys Pro Asp Ser Asp  
 100 105 110  
 Glu Pro Lys Gly Ser Gly Phe Asp Pro Glu Ser Leu Glu Arg Gly Ala  
 115 120 125

170

SUBSTITUTE SHEET (RULE 26)

Lys Ala Leu Arg Glu Ile Asn Ser Ser Pro His Ser Lys Lys Val Phe  
 130 135 140  
 Asp Leu Met Arg Lys Gln Glu Lys Thr Arg Leu Ala Glu Leu Ala Ala  
 145 150 155 160  
 Glu Lys Glu His Asn Glu Ala Ile Gln Ala Ser Lys Asp Ile Glu Arg  
 165 170 175  
 Gln Arg Lys Leu Ala Glu Asp Gln Arg Asn Leu Val Gln Gln Glu Ala  
 180 185 190  
 Gln Ala Lys Ala Gln Asn Leu Arg Tyr Glu Asp Glu Leu Ala Arg Lys  
 195 200 205  
 Arg Met Gln Thr Asp Asn Glu Ala Gln Arg Arg His Asn Ala Glu Leu  
 210 215 220  
 Val Ser Met Gln Glu Ala Ser Ser Ile Arg Lys Glu Lys Ala Arg Ile  
 225 230 235 240  
 Ala Thr Glu Glu Gln Ile Gln Ala Gln Gln Arg Glu Thr Glu Lys Glu  
 245 250 255  
 Arg Ala Glu Leu Glu Arg Glu Thr Ile Arg Val Lys Ala Met Ala Glu  
 260 265 270  
 Ala Glu Gly Arg Ala His Glu Ala Lys Leu Thr Glu Glu Gln Asn Arg  
 275 280 285  
 Arg Met Leu Leu Asp Lys Ile Asn Gly Glu Arg Glu Lys Trp Leu Ala  
 290 295 300  
 Ala Ile Asn Thr Thr Phe Ser His Ile Glu Gly Gly Val Arg Thr Leu  
 305 310 315 320  
 Leu Thr Asp Arg Ser Lys Leu Ile Met Thr Val Gly Gly Val Thr Ala  
 325 330 335  
 Leu Ala Ala Gly Val Tyr Thr Thr Arg Glu Gly Ala Arg Val Thr Trp  
 340 345 350  
 Gly Tyr Ile Asn Arg Ile Leu Gly Gln Pro Ser Leu Ile Arg Glu Ser  
 355 360 365  
 Ser Met Gly Arg Phe Pro Trp Ala Gly Ser Val Ser Gln Phe Lys Asn  
 370 375 380  
 Lys Leu Ser Thr Ala Ala Gly Ala Ala Ser Ala Glu Gly Glu Lys  
 385 390 395 400  
 Pro Leu Glu Asn Val Ile Leu His Arg Ser Leu Lys Thr Arg Ile Glu  
 405 410 415  
 Arg Leu Ala Arg Ala Thr Ala Asn Thr Lys Ser His Lys Ala Pro Phe  
 420 425 430  
 Arg Asn Met Met Phe Tyr Gly Pro Pro Gly Thr Gly Lys Thr Met Val  
 435 440 445  
 Ala Arg Glu Ile Ala Arg Lys Ser Gly Leu Asp Tyr Ala Met Met Thr  
 450 455 460  
 Gly Gly Asp Val Ala Pro Leu Gly Ala Gln Ala Val Thr Lys Ile His  
 465 470 475 480  
 Glu Ile Phe Asp Trp Ala Lys Lys Ser Asn Lys Gly Leu Leu Phe  
 485 490 495  
 Ile Asp Glu Ala Asp Ala Phe Leu Cys Glu Arg Asn Ser Thr Tyr Met  
 500 505 510  
 Ser Glu Ala Gln Arg Ser Ala Leu Asn Ala Leu Leu Phe Arg Thr Gly  
 515 520 525  
 Asp Gln Ser Arg Asp Ile Val Leu Val Leu Ala Thr Asn Arg Pro Gly  
 530 535 540  
 Asp Leu Asp Ser Ala Val Thr Asp Arg Ile Asp Glu Val Ile Glu Phe  
 545 550 555 560  
 Pro Leu Pro Gly Glu Glu Arg Phe Lys Leu Leu Lys Leu Tyr Leu  
 565 570 575  
 Asn Lys Tyr Leu Met Gly Asp Asp Lys Lys Gly Glu Lys Asp Ser Asn  
 580 585 590  
 Leu Lys Trp Ser Asn Leu Phe Lys Lys Lys Ser Gln Lys Ile Thr  
 595 600 605  
 Ile Glu Gly Asp Leu Thr Thr Asp Gln Val Ile Lys Glu Ala Ala Lys Lys

Thr Glu Gly Phe Ser Gly Arg Glu Ile Ala Lys Leu Val Ala Gly Val  
 610 615 620  
 625 630 635 640  
 Gln Ala Ala Val Tyr Gly Arg Gln Asp Cys Val Leu Asp Ser Gln Leu  
 645 650 655  
 Phe Glu Glu Ile Val Asp Tyr Lys Ile Glu Glu His His Gln Arg Ile  
 660 665 670  
 Arg Leu Ala Thr Glu Gly Gln Ser Phe Pro  
 675 680  
 <210> 348  
 <211> 953  
 <212> DNA  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..953:Ceres Seq. ID 1975983  
 <400>348  
 aaaaagata aacaaasgc tgaaggagg aaaccagagg aaggagagca gacaccacc  
 cacaagcat ctacgtctt ccttaataca tggcagacaa gggcgcccg ttgcgaagt  
 ttggaagc tgaatgcgt tgggtggg tgggtgctc ctcttcaat ctacgcaat  
 tcatgatat atccgagat aatcgctg agatctcaa gtacctctc aaagagggtg  
 ttgtgtgc caaaagat ttcaattac cacaaatcc ttgattgag agtggtccaa  
 atctgaagt tatcaagt atcgagagt tcaaatctaa ggaatatgtg agagagacct  
 ttgttgat gcatctac tggttctca caaatgaag tattgacttt cttaggactt  
 accttaatt ccatctgag attgttctg ctactctgaa gaagcaacag aagcctttg  
 gtgacactt tggagtggt ggtgacctc ccctggccc tctctggt gatggagaa  
 gggaggttgg tgacagagt gataccgtg gaggctctaa atcagggtga gattaggtg  
 acaagctgg agcacctgt gattaccag ctggcttcag ggtggagct agtgagcaa  
 ggcaaggtt tggctgga gctggtgt ttggtgtg tctgtgcca gctctggat  
 ctgatatcc ttgaaagga cttcttgt ttctttgtt ctatttaag gtacataac  
 acctatgga gaacagagt gcttttga acctgttc ttctcttaa acaattcac  
 aatagattg ttttttgc ttgaattt atgaagaat gttgacgt lacatgaac  
 tttttacat gtctctctt ttaaatctc ctattagtt tttttgaag ctt  
 <210> 349  
 <211> 214  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..214:Ceres Seq. ID 1975984  
 <400>349  
 Met Ala Asp Lys Gly Arg Pro Leu Pro Lys Phe Gly Arg Ala Glu Met  
 1 5 10 15  
 Arg Cys Gly Gly Gly Val Ala Ser Phe Phe Asn Leu Ile Ala Ile Met  
 20 25 30  
 Ile Ile Ser Glu Thr Asn Arg Arg Glu Ile Ser Lys Tyr Leu Phe Lys  
 35 40 45  
 Glu Gly Val Leu Phe Ala Lys Lys Asp Phe Asn Leu Pro Gln His Pro  
 50 55 60  
 Leu Ile Glu Ser Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln Ser  
 65 70 75  
 Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His Tyr  
 85 90 95



Tyr Trp Phe Leu Thr Asn Glu Gly Ile Asp Phe Leu Arg Thr Tyr Leu  
100 110  
Asn Leu Pro Ser Glu Ile Val Pro Ala Thr Leu Lys Lys Gln Gln Lys  
115 125  
Pro Leu Gly Arg Pro Phe Gly Gly Gly Asp Arg Pro Arg Gly Pro  
130 140  
Pro Arg Gly Asp Gly Glu Arg Arg Phe Gly Asp Arg Asp Gly Tyr Arg  
145 150 155 160  
Gly Gly Pro Lys Ser Gly Gly Glu Tyr Gly Asp Lys Ala Gly Ala Pro  
165 170 175  
Ala Asp Tyr Gln Pro Gly Phe Arg Gly Gly Ala Ser Gly Ala Arg Gln  
180 185 190  
Gly Phe Gly Arg Gly Ala Gly Gly Phe Gly Gly Ala Gly Pro Ala  
195 200 205  
Ala Gly Ser Asp Leu Pro  
210

<210> 350  
<211> 199  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..199:Ceres Seq. ID 1975985

<400>350  
Met Arg Cys Gly Gly Val Ala Ser Phe Phe Asn Leu Ile Ala Ile  
1 5 10 15  
Met Ile Ser Glu Thr Asn Arg Arg Glu Ile Ser Lys Tyr Leu Phe  
20 25 30  
Lys Glu Gly Val Leu Phe Ala Lys Lys Asp Phe Asn Leu Pro Gln His  
35 40 45  
Pro Leu Ile Glu Ser Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln  
50 55 60  
Ser Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His  
65 70 75 80  
Tyr Tyr Trp Phe Leu Thr Asn Glu Gly Ile Asp Phe Leu Arg Thr Tyr  
85 90 95  
Leu Asn Leu Pro Ser Glu Ile Val Pro Ala Thr Leu Lys Lys Gln Gln  
100 105 110  
Lys Pro Leu Gly Arg Pro Phe Gly Gly Gly Asp Arg Pro Arg Gly  
115 120 125  
Pro Pro Arg Gly Asp Gly Glu Arg Arg Phe Gly Asp Arg Asp Gly Tyr  
130 135 140  
Arg Gly Gly Pro Lys Ser Gly Gly Glu Tyr Gly Asp Lys Ala Gly Ala  
145 150 155 160  
Pro Ala Asp Tyr Gln Pro Gly Phe Arg Gly Gly Ala Ser Gly Ala Arg  
165 170 175  
Gln Gly Phe Gly Arg Gly Ala Gly Gly Phe Gly Gly Gly Ala Gly Pro  
180 185 190  
Ala Ala Gly Ser Asp Leu Pro  
195

<210> 351  
<211> 183  
<212> PRT  
<213> Arabidopsis thaliana

<220>

173

SUBSTITUTE SHEET (RULE 26)

<223> any n or Xaa = unknown

<223> LOCATION: 1..183:Ceres Seq. ID 1975986

<400>351  
Met Ile Ile Ser Glu Thr Asn Arg Arg Glu Ile Ser Lys Tyr Leu Phe  
1 5 10 15  
Lys Glu Gly Val Leu Phe Ala Lys Lys Asp Phe Asn Leu Pro Gln His  
20 25 30  
Pro Leu Ile Glu Ser Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln  
35 40 45  
Ser Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His  
50 55 60  
Tyr Tyr Trp Phe Leu Thr Asn Glu Gly Ile Asp Phe Leu Arg Thr Tyr  
65 70 75 80  
Leu Asn Leu Pro Ser Glu Ile Val Pro Ala Thr Leu Lys Lys Gln Gln  
85 90 95  
Lys Pro Leu Gly Arg Pro Phe Gly Gly Gly Asp Arg Pro Arg Gly  
100 105 110  
Pro Pro Arg Gly Asp Gly Glu Arg Arg Phe Gly Asp Arg Asp Gly Tyr  
115 120 125  
Arg Gly Gly Pro Lys Ser Gly Gly Glu Tyr Gly Asp Lys Ala Gly Ala  
130 135 140  
Pro Ala Asp Tyr Gln Pro Gly Phe Arg Gly Gly Ala Ser Gly Ala Arg  
145 150 155 160  
Gln Gly Phe Gly Arg Gly Ala Gly Gly Phe Gly Gly Ala Gly Pro  
165 170 175  
Ala Ala Gly Ser Asp Leu Pro  
180

<210> 352  
<211> 1027  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..1027:Ceres Seq. ID 1976019

<400>352  
atcgaataa ggggttcgcy ttaggagaag aagttaangc aaacacata caaacgcagt  
cacctctct gtgcctcct tcttcaatc catcgcaatc atgacatc cgcagactaa  
tcgccgtgag atgcctcaagt acctctcaa agagggtgt ttgttgcca aaagagattt  
caattacca caacatcctt tgattgagag tgttccaat ctgcaagtta tcaagtgat  
gcaggtttc aatctaagg aatatgag agagacctt gcttgatgc attactagt  
gttctcaca aatgaagta ttgactttc taggacttac cttaactcc catctgagat  
tgttctcgt acctgaaga agcaacaga gctcttggc cgacctttg gaggtgtgg  
tgacctccc cgtggccctc ctctgtga tggagaggg aggttggg acagagatgg  
atccgtgga ggtctaaat caggtgaga gtatgtgac aagctggag cactgtga  
ttaccagct gcttcaggg gtgagctag tggagaagg caaggtttg gtcgtggagc  
tgggttttt ggtgtgggt cgtgtcagc tgttgatct gattacctt gaaagggaga  
cctacagct cagcacaagc caagccttt atccctac acaatgag cggacatga  
gcttcacaa tcaacatcc catctcagt caccaatc tagctatg agacagaa  
tggcctcta aacctggag ctctcttta attgcaggt gtaacaatg agtaagag  
tgatgggca atcagttt gcaggtaga atcagagc ttattgtac aataacaa  
tattacaa catgcttat atagatgc tacttttt ctctgtta acgataacg  
cagtaacct attgtacc tctctac atataatg acccaatta attatcata  
ggctttc

<210> 353

174

SUBSTITUTE SHEET (RULE 26)

<211> 183  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..183:Ceres Seq. ID 1976020

<400>353  
 Met Ile Ile Ser Glu Thr Asn Arg Arg Glu Ile Ser Lys Tyr Leu Phe  
 1 5 10 15  
 Lys Glu Gly Val Leu Phe Ala Lys Lys Asp Phe Asn Leu Pro Gln His  
 20 25 30  
 Pro Leu Ile Glu Ser Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln  
 35 40 45  
 Ser Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His  
 50 55 60  
 Tyr Tyr Trp Phe Leu Thr Asn Glu Gly Ile Asp Phe Leu Arg Thr Tyr  
 65 70 75  
 Leu Asn Leu Pro Ser Glu Ile Val Pro Ala Thr Leu Lys Lys Gln Gln  
 85 90 95  
 Lys Pro Leu Gly Arg Pro Phe Gly Gly Gly Asp Arg Pro Arg Gly  
 100 105 110  
 Pro Pro Arg Gly Asp Gly Glu Arg Arg Phe Gly Asp Arg Asp Gly Tyr  
 115 120 125  
 Arg Gly Gly Pro Lys Ser Gly Gly Glu Tyr Gly Asp Lys Ala Gly Ala  
 130 135 140  
 Pro Ala Asp Tyr Gln Pro Gly Phe Arg Gly Gly Ala Ser Gly Ala Arg  
 145 150 155  
 Gln Gly Phe Gly Arg Gly Ala Gly Gly Phe Gly Gly Ala Gly Pro  
 165 170 175  
 Ala Ala Gly Ser Asp Leu Pro  
 180

<210> 354  
 <211> 137  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..137:Ceres Seq. ID 1976021

<400>354  
 Met Gln Ser Phe Lys Ser Lys Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp  
 1 5 10 15  
 Met His Tyr Tyr Trp Phe Leu Thr Asn Glu Gly Ile Asp Phe Leu Arg  
 20 25 30  
 Thr Tyr Leu Asn Leu Pro Ser Glu Ile Val Pro Ala Thr Leu Lys Lys  
 35 40 45  
 Gln Gln Lys Pro Leu Gly Arg Pro Phe Gly Gly Gly Asp Arg Pro  
 50 55 60  
 Arg Gly Pro Pro Arg Gly Asp Gly Glu Arg Arg Phe Gly Asp Arg Asp  
 65 70 75  
 Gly Tyr Arg Gly Gly Pro Lys Ser Gly Gly Glu Tyr Gly Asp Lys Ala  
 85 90 95  
 Gly Ala Pro Ala Asp Tyr Gln Pro Gly Phe Arg Gly Gly Ala Ser Gly  
 100 105 110  
 Ala Arg Gln Gly Phe Gly Arg Gly Ala Gly Phe Gly Gly Gly Ala

115 120 125  
 Gly Pro Ala Ala Gly Ser Asp Leu Pro  
 130

<210> 355  
 <211> 121  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..121:Ceres Seq. ID 1976022

<400>355  
 Met His Tyr Tyr Trp Phe Leu Thr Asn Glu Gly Ile Asp Phe Leu Arg  
 1 5 10 15  
 Thr Tyr Leu Asn Leu Pro Ser Glu Ile Val Pro Ala Thr Leu Lys Lys  
 20 25 30  
 Gln Gln Lys Pro Leu Gly Arg Pro Phe Gly Gly Gly Asp Arg Pro  
 35 40 45  
 Arg Gly Pro Pro Arg Gly Asp Gly Glu Arg Arg Phe Gly Asp Arg Asp  
 50 55 60  
 Gly Tyr Arg Gly Gly Pro Lys Ser Gly Gly Glu Tyr Gly Asp Lys Ala  
 65 70 75  
 Gly Ala Pro Ala Asp Tyr Gln Pro Gly Phe Arg Gly Gly Ala Ser Gly  
 85 90 95  
 Ala Arg Gln Gly Phe Gly Arg Gly Ala Gly Gly Phe Gly Gly Gly Ala  
 100 105 110  
 Gly Pro Ala Ala Gly Ser Asp Leu Pro  
 115 120

<210> 356  
 <211> 478  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..478:Ceres Seq. ID 1976673

<400>356  
 aaattcaatc tcttcaatt tcttcttc tctctgcga cgcgttaatg cgcgcttcgt  
 tctcaatcac gacttcact tcttctatc caccattcaa atctcaaac aaactaac  
 60  
 caccacaaa tctactctt ccttcccaa ctatcccca aggcgaaga atgattctg  
 120  
 ctatcgaaa aatggcggtc gaagaattt cttaaccgc tcttctact tctctgagc  
 180  
 ttgcttctgt gatagccgc tgccttgctt acccaaacac gctcttcttc agttctggat  
 240  
 acaatgtgca agtcttctgt gaagataacg agtcagagga gaggcttggtg aatcgattta  
 300  
 ggagagaact gatgaact ggtgtatatac aggaatgaa gaggaaga tactttgaga  
 360  
 ataacaaga tgagaagaaa cgtaggactc gtagctgc taagcgtaat aagaaaag  
 420  
 478

<210> 357  
 <211> 158  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..158:Ceres Seq. ID 1976674

<400>357  
Ile Gln Ser Leu Pro Ile Ser Leu Leu Leu Pro Arg Asp Ala Leu Met  
1 5 10 15  
Ala Ala Ser Phe Ser Leu Thr Ser Phe Ile Ser Phe Ile Ser Pro Phe  
20 25 30  
Lys Ser Gln Thr Lys Pro Thr Pro Pro Pro Asn Leu Thr Leu Pro Ser  
35 40 45  
Pro Thr Ile Ser Gln Arg Arg Asn Asp Leu Ala Ile Gln Ser Met  
50 55 60  
Ala Val Gln Glu Ser Ser Thr Ala Ser Ser Ser Leu Ser Ser Glu Leu  
65 70 75 80  
Ala Ser Val Ile Cys Pro Ser Leu Ala Tyr Ser Asn Thr Leu Phe Phe  
85 90 95  
Ser Ser Gly Tyr Asn Val Gln Val Phe Val Glu Asp Asn Glu Ser Glu  
100 105 110  
Glu Arg Leu Val Asn Arg Phe Arg Arg Glu Val Met Arg Thr Gly Val  
115 120 125  
Ile Gln Glu Cys Lys Arg Arg Arg Tyr Phe Glu Asn Lys Gln Asp Glu  
130 135 140  
Lys Lys Arg Arg Thr Arg Asp Ala Ala Lys Arg Asn Lys Lys  
145 150 155

&lt;210&gt; 358

&lt;211&gt; 143

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..143:Ceres Seq. ID 1976675

<400>358  
Met Ala Ala Ser Phe Ser Leu Thr Ser Phe Ile Ser Phe Ile Ser Pro  
1 5 10 15  
Phe Lys Ser Gln Thr Lys Pro Thr Pro Pro Pro Asn Leu Thr Leu Pro  
20 25 30  
Ser Pro Thr Ile Ser Gln Arg Arg Arg Asn Asp Leu Ala Ile Glu Ser  
35 40 45  
Met Ala Val Glu Glu Ser Ser Thr Ala Ser Ser Leu Ser Ser Glu  
50 55 60  
Leu Ala Ser Val Ile Cys Pro Ser Leu Ala Tyr Ser Asn Thr Leu Phe  
65 70 75 80  
Phe Ser Ser Gly Tyr Asn Val Gln Val Phe Val Glu Asp Asn Glu Ser  
85 90 95  
Glu Glu Arg Leu Val Asn Arg Phe Arg Arg Glu Val Met Arg Thr Gly  
100 105 110  
Val Ile Gln Glu Cys Lys Arg Arg Arg Tyr Phe Glu Asn Lys Gln Asp  
115 120 125  
Glu Lys Lys Arg Arg Thr Arg Asp Ala Ala Lys Arg Asn Lys Lys  
130 135 140 145

&lt;210&gt; 359

&lt;211&gt; 95

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..95:Ceres Seq. ID 1976676

<400>359  
Met Ala Val Glu Glu Ser Ser Thr Ala Ser Ser Leu Ser Ser Glu  
1 5 10 15  
Leu Ala Ser Val Ile Cys Pro Ser Leu Ala Tyr Ser Asn Thr Leu Phe  
20 25 30  
Phe Ser Ser Gly Tyr Asn Val Gln Val Phe Val Glu Asp Asn Glu Ser  
35 40 45  
Glu Glu Arg Leu Val Asn Arg Phe Arg Arg Glu Val Met Arg Thr Gly  
50 55 60  
Val Ile Gln Glu Cys Lys Arg Arg Tyr Phe Glu Asn Lys Gln Asp  
65 70 75 80  
Glu Lys Lys Arg Arg Thr Arg Asp Ala Ala Lys Arg Asn Lys Lys  
85 90 95

&lt;210&gt; 360

&lt;211&gt; 1076

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..1076:Ceres Seq. ID 2025186

&lt;400&gt;360

aatcattgc agaagaaaaa taataaaaaa aaacagagaa aaggagatat ggaatcaatc  
60  
ggagtcctta tgaatgccc catgtcctcc tactcagaa acgagcttga gaagcgcttc  
120  
aacctcttc gctctggag tctccggag aaatccgtcc tctagaaac tcatgggaa  
180  
tccatccgc ccgtcgttgg gaatgcttt ccgcggccg atgtccagct catcagcat  
240  
ctgccaac ttgagattgt atccagcttc agctcgccg tcgacaagat cgatttggg  
300  
aaatcaaa aaagaaggat ccgcgtcac aaacccccc acgttctcac caaagacgc  
360  
gaagatctcg catcgggct tatctggct cctctcgac ggctgtgtga gtgcgctgc  
420  
tatgtagga gcggaaaatg gaagcaaggt gaattccaa tcaatccaa gtttagtga  
480  
aaatccgtgg ggaatcattg tctagtaga attggactg ccacgcgaaa gaggctgaa  
540  
gcctttagct gcccaatcaa ttactactca agaaccatta agcctgatgt cgcctacaag  
600  
tattatcga cggtyggtga ccttgcctaa aactcagaca tccctcgt cgcgtgccg  
660  
ttgaccgagc agaccagaca cattgtggac cggcaggtca tggatgcatt aggaactaag  
720  
ggcgtctca taacattgg ccgtggacca catgttgatg agcaagact tataaagct  
780  
ctaacagaag ccgcctagg tggggtgcc ctgtgatgt ttgagcaga gcctcacgtg  
840  
cccgaggagc tcttggcct tgagaatga gtctctccc ctcaatgg gagtggcact  
900  
gtggaaacac ggaatgccat ggcgatctt gtctgggta acttgaagc gcactttct  
960  
gggaatcac tctgactcc ggtgcttga gtcagtgc acatttggga ttgtgttca  
1020  
ctctattga aagggtttta tgttaacgaa ctcatgaaa gggtaatttc tcttcc  
1076

&lt;210&gt; 361

&lt;211&gt; 329

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..329:Ceres Seq. ID 2025187

&lt;400&gt;361

Asn His Leu Gln Lys Lys Asn Lys Lys Lys Lys Gln Arg Lys Gly Asp  
1 5 10 15  
Met Glu Ser Ile Gly Val Leu Met Met Cys Pro Met Ser Ser Tyr Leu  
20 25 30

Glu Asn Glu Leu Glu Lys Arg Phe Asn Leu Leu Arg Phe Thr Thr Ser  
35 40 45  
Pro Glu Lys Ser Val Leu Leu Glu Thr His Arg Asn Ser Ile Arg Ala  
50 55 60  
Val Val Gly Asn Ala Ser Ala Gly Ala Asp Ala Gln Leu Ile Ser Asp  
65 70 75  
Leu Pro Asn Leu Glu Ile Val Ser Ser Phe Ser Val Gly Leu Asp Lys  
85 90 95  
Ile Asp Leu Gly Lys Cys Lys Glu Lys Gly Ile Arg Val Thr Asn Thr  
100 105 110  
Pro Asp Val Leu Thr Glu Asp Val Ala Asp Leu Ala Ile Gly Leu Ile  
115 120 125  
Leu Ala Leu Leu Arg Arg Leu Cys Glu Cys Asp Arg Tyr Val Arg Ser  
130 135 140  
Gly Lys Trp Lys Gln Gly Glu Phe Gln Leu Thr Lys Phe Ser Gly  
145 150 155  
Lys Ser Val Gly Ile Ile Gly Leu Gly Arg Ile Gly Thr Ala Ile Ala  
165 170 175  
Lys Arg Ala Glu Ala Phe Ser Cys Pro Ile Asn Tyr Tyr Ser Arg Thr  
180 185 190  
Ile Lys Pro Asp Val Ala Tyr Lys Tyr Tyr Pro Thr Val Val Asp Leu  
195 200 205  
Ala Gln Asn Ser Asp Ile Leu Val Val Ala Cys Pro Leu Thr Glu Gln  
210 215 220  
Thr Arg His Ile Val Asp Arg Gln Val Met Asp Ala Leu Gly Ala Lys  
225 230 235  
Gly Val Leu Ile Asn Ile Gly Arg Gly Pro His Val Asp Glu Glu  
245 250 255  
Leu Ile Lys Ala Leu Thr Glu Gly Arg Leu Gly Gly Ala Ala Leu Asp  
260 265 270  
Val Phe Glu Gln Glu Pro His Val Pro Glu Glu Leu Phe Gly Leu Glu  
275 280 285  
Asn Val Val Leu Leu Pro His Val Pro Glu Glu Leu Phe Gly Leu Glu  
290 295 300  
Asn Ala Met Ala Asp Leu Val Val Gly Asn Leu Glu Ala His Phe Ser  
305 310 315  
Gly Lys Ser Leu Leu Thr Pro Val Val  
325

&lt;210&gt; 362

&lt;211&gt; 313

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1...313:Ceres Seq. ID 2025188

&lt;400&gt;362

Met Glu Ser Ile Gly Val Leu Met Met Cys Pro Met Ser Ser Tyr Leu  
1 5 10 15  
Glu Asn Glu Leu Glu Lys Arg Phe Asn Leu Leu Arg Phe Thr Thr Ser  
20 25 30  
Pro Glu Lys Ser Val Leu Leu Glu Thr His Arg Asn Ser Ile Arg Ala  
35 40 45  
Val Val Gly Asn Ala Ser Ala Gly Ala Asp Ala Gln Leu Ile Ser Asp  
50 55 60  
Leu Pro Asn Leu Glu Ile Val Ser Ser Phe Ser Val Gly Leu Asp Lys  
65 70 75  
Ile Asp Leu Gly Lys Cys Lys Glu Lys Gly Ile Arg Val Thr Asn Thr

Pro Asp Val Leu Thr Glu Asp Val Ala Asp Leu Ala Ile Gly Leu Ile  
85 90 95  
Leu Ala Leu Leu Arg Arg Leu Cys Glu Cys Asp Arg Tyr Val Arg Ser  
100 105 110  
Gly Lys Trp Lys Gln Gly Glu Phe Gln Leu Thr Thr Lys Phe Ser Gly  
115 120 125  
Lys Ser Val Gly Ile Ile Gly Leu Gly Arg Ile Gly Thr Ala Ile Ala  
130 135 140  
Lys Arg Ala Gln Ala Phe Ser Cys Pro Ile Asn Tyr Tyr Ser Arg Thr  
145 150 155  
Ile Lys Pro Asp Val Ala Tyr Lys Tyr Tyr Pro Thr Val Val Asp Leu  
165 170 175  
Ala Gln Asn Ser Asp Ile Leu Val Val Ala Cys Pro Leu Thr Glu Gln  
180 185 190  
Thr Arg His Ile Val Asp Arg Gln Val Met Asp Ala Leu Gly Ala Lys  
195 200 205  
Gly Val Leu Ile Asn Ile Gly Arg Gly Pro His Val Asp Glu Gln Glu  
210 215 220  
Leu Ile Lys Ala Leu Thr Glu Gly Arg Leu Gly Gly Ala Ala Leu Asp  
225 230 235  
Val Phe Glu Gln Glu Pro His Val Pro Glu Glu Leu Phe Gly Leu Glu  
245 250 255  
Asn Val Val Leu Leu Pro His Val Gly Ser Gly Thr Val Glu Thr Arg  
260 265 270  
Asn Ala Met Ala Asp Leu Val Val Gly Asn Leu Glu Ala His Phe Ser  
275 280 285  
Gly Lys Ser Leu Leu Thr Pro Val Val  
290 295 300  
305 310

&lt;210&gt; 363

&lt;211&gt; 306

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1...306:Ceres Seq. ID 2025189

&lt;400&gt;363

Met Met Cys Pro Met Ser Ser Tyr Leu Glu Asn Glu Leu Glu Lys Arg  
1 5 10 15  
Phe Asn Leu Leu Arg Phe Thr Thr Ser Pro Glu Lys Ser Val Leu Leu  
20 25 30  
Glu Thr His Arg Asn Ser Ile Arg Ala Val Val Gly Asn Ala Ser Ala  
35 40 45  
Gly Ala Asp Ala Gln Leu Ile Ser Asp Leu Pro Asn Leu Glu Ile Val  
50 55 60  
Ser Ser Phe Ser Val Gly Leu Asp Lys Ile Asp Leu Gly Lys Cys Lys  
65 70 75  
Glu Lys Gly Ile Arg Val Thr Asn Thr Pro Asp Val Leu Thr Glu Asp  
85 90 95  
Val Ala Asp Leu Ala Ile Gly Leu Ile Leu Ala Leu Leu Arg Arg Leu  
100 105 110  
Cys Glu Cys Asp Arg Tyr Val Arg Ser Gly Lys Trp Lys Gln Gly Glu  
115 120 125  
Phe Gln Leu Thr Thr Lys Phe Ser Gly Lys Ser Val Gly Ile Ile Gly  
130 135 140  
Leu Gly Arg Ile Gly Thr Ala Ile Ala Lys Arg Ala Glu Ala Phe Ser  
145 150 155 160

Cys Pro Ile Asn Tyr Tyr Ser Arg Thr Ile Lys Pro Asp Val Ala Tyr  
165 170 175  
Lys Tyr Tyr Pro Thr Val Val Asp Leu Ala Gln Asn Ser Asp Ile Leu  
180 185 190  
Val Val Ala Cys Pro Leu Thr Glu Gln Thr Arg His Ile Val Asp Arg  
195 200 205  
Gln Val Met Asp Ala Leu Gly Ala Lys Gly Val Leu Ile Asn Ile Gly  
210 215 220  
Arg Gly Pro His Val Asp Glu Gln Leu Ile Lys Ala Leu Thr Glu  
225 230 235  
Gly Arg Leu Gly Gly Ala Ala Leu Asp Val Phe Glu Gln Glu Pro His  
245 250 255  
Val Pro Glu Glu Leu Phe Gly Leu Glu Asn Val Val Leu Leu Pro His  
260 265 270  
Val Gly Ser Gly Thr Val Glu Thr Arg Asn Ala Met Ala Asp Leu Val  
275 280 285  
Val Gly Asn Leu Glu Ala His Phe Ser Gly Lys Ser Leu Leu Thr Pro  
290 295 300  
Val Val  
305

&lt;210&gt; 364

&lt;211&gt; 555

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..555:Ceres Seq. ID 2025372

&lt;400&gt;364

aaacacact ttcaactgca caagataag gaaccatgt ctgtgtcagc gatcttgg  
accggaatg tacccttgc tgcctccg gttccgcc aatttcagt tccaaatg  
ggtaaggag gaggattag gaggtaga gaggtagt cagggccaca gaagaaatg  
acggtcacc acaggagac gaggcgaag aagactcagc ctgggacat taagaaag  
ctacgtgtt acctcctct tctcctctt ccggcggat ggaatcgtt caetctcgt  
tctgacgag gtggtcgtc cactctcgc ggaatttgg ttccaggcc tgcctagt  
gtatgagta tctcgtgatt tggttgta cagtttgg aactttgtt gtcgttact  
gtgtttcag ttctctgatg ttgttggg ggtttgtg ttgttgatg gtgaatgat  
ggaccaaat tgcgattat aagcttcaa ccttctcct ctgttttga gtttaaggt  
ctcaactta tagt

&lt;210&gt; 365

&lt;211&gt; 118

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..118:Ceres Seq. ID 2025373

&lt;400&gt;365

Lys Thr His Phe His Leu His Lys Asp Lys Glu Thr Met Ser Val Ser  
1 5 10 15  
Ala Ile Phe Gly Thr Gly Ile Val Thr Val Ala Ala Ser Pro Val Leu  
20 25 30  
Arg Gln Phe Gln Val Pro Lys Leu Gly Asn Gly Gly Leu Gly Met  
35 40 45  
Val Ile Glu Cys Ser Arg Pro Gln Lys Lys Ser Thr Ala His His

Arg Lys Thr Arg Pro Lys Lys Thr Gln Pro Trp Asp Ile Lys Arg Lys  
55 60  
65 70 75  
Pro Thr Val Tyr Ala Pro Leu Pro Pro Leu Ala Glu Trp Ser Pro  
85 90 95  
Phe Thr Leu Ala Ser Asp Asp Gly Gly Ala Ala Thr Ala Ala Gly Asp  
100 105 110  
Leu Val Ser Gly Ala Ala  
115

&lt;210&gt; 366

&lt;211&gt; 106

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..106:Ceres Seq. ID 2025374

&lt;400&gt;366

Met Ser Val Ser Ala Ile Phe Gly Thr Gly Ile Val Thr Val Ala Ala  
1 5 10 15  
Ser Pro Val Leu Arg Gln Phe Gln Val Pro Lys Leu Gly Asn Gly Gly  
20 25 30  
Gly Leu Gly Met Val Ile Glu Cys Ser Ser Arg Pro Gln Lys Lys Ser  
35 40 45  
Thr Ala His His Arg Lys Thr Arg Pro Lys Lys Thr Gln Pro Trp Asp  
50 55 60  
Ile Lys Arg Lys Pro Thr Val Tyr Ala Pro Leu Pro Pro Leu Pro Ala  
65 70 75  
Glu Trp Ser Pro Phe Thr Leu Ala Ser Asp Asp Gly Gly Ala Thr  
85 90 95  
Ala Ala Gly Asp Leu Val Ser Gly Ala Ala  
100 105

&lt;210&gt; 367

&lt;211&gt; 71

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..71:Ceres Seq. ID 2025375

&lt;400&gt;367

Met Val Ile Glu Cys Ser Ser Arg Pro Gln Lys Lys Ser Thr Ala His  
1 5 10 15  
His Arg Lys Thr Arg Pro Lys Lys Thr Gln Pro Trp Asp Ile Lys Arg  
20 25 30  
Lys Pro Thr Val Tyr Ala Pro Leu Pro Pro Leu Pro Ala Glu Trp Ser  
35 40 45  
Pro Phe Thr Leu Ala Ser Asp Asp Gly Gly Ala Thr Ala Ala Gly  
50 55 60  
Asp Leu Val Ser Gly Ala Ala  
65 70

&lt;210&gt; 368

&lt;211&gt; 631

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..631:Ceres Seq. ID 2025471

&lt;400&gt;368

atgctgcgaa agttctctaa accgccaagt gtcagtcact tatccacag taatcccaatt 60  
 attctcaag cctccggttt attacogtct aaccgaataa gaaaactc tcttcggaat 120  
 ctgaaagcc ataaacagc agaaaataat ctaaggatc ggaatttta taatcggaat 180  
 aatagaatg tatgcagtt ttcggtgaag ttgtaagagg agaagagt taccgggtga 240  
 agcggcgag atagacttt aagatctgcc gtcaagaat cgtctcgtc tccacagat 300  
 gcgtctcgt cgtctcgtt gaggagctt aaaggatc tgaatcgag tegtattggt 360  
 gcggcgca gtgagaggt gagcaagct gaggatctt tgaggcgtt gatgttctg 420  
 agctttggg gatctgttta gatctgtttg agaaaataat agatgagaaa acgaaccaa 480  
 aacgtttcg ttgtgtttt tgattttgg attttgtt tctctgtgaa tagtttttt 540  
 tttttttc gttctcttg tacttttgt tttatgtaaa tcatgtataa aatgaaagt 600  
 aaatcaatga agataatgtt ttgtcttcac c 631

&lt;210&gt; 369

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..146:Ceres Seq. ID 2025472

&lt;400&gt;369

Met Ser Arg Lys Phe Leu Lys Pro Pro Arg Val Ser His Leu Ser Asn 1  
 1  
 Ser Asn Pro Ile Ile Arg Gln Ala Ser Gly Leu Leu Pro Ser Lys Pro 10  
 20 25 30  
 Lys Arg Lys Asn Ser Leu Arg Asn Leu Lys Ser His Lys Ser Ala Glu 35  
 40 45  
 Lys Asn Leu Lys Asp Arg Lys Phe Tyr Asn Arg Met Asn Arg Val Ile 50  
 55 60  
 Ser Gln Phe Ser Gly Lys Leu Met Lys Glu Lys Ser Val Thr Gly Val 65  
 70 75 80  
 Ser Gly Gly Asp Thr Thr Leu Arg Ser Ala Val Lys Glu Ser Val Ser 85  
 90 95  
 Ser Pro Gln Ser Ala Ser Ser Ser Ser Val Arg Arg Leu Lys Gly 100  
 105 110  
 Asp Leu Glu Ser Ser Arg Phe Gly Ala Ala Ala Ser Glu Arg Leu Arg 115  
 120 125  
 Gln Ala Glu Glu Ser Leu Arg Thr Val Met Phe Leu Ser Cys Trp Gly 130  
 135 140  
 Ser Cys 145

&lt;210&gt; 370

&lt;211&gt; 87

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..87:Ceres Seq. ID 2025473

&lt;400&gt;370

Met Asn Arg Val Ile Ser Gln Phe Ser Gly Lys Leu Met Lys Glu Lys 1  
 5 10 15  
 Ser Val Thr Gly Val Ser Gly Gly Asp Thr Thr Leu Arg Ser Ala Val 20  
 25 30  
 Lys Glu Ser Val Ser Ser Pro Gln Ser Ala Ser Ser Ser Ser Val 35  
 40 45  
 Arg Arg Leu Lys Gly Asp Leu Glu Ser Ser Arg Phe Gly Ala Ala Ala 50  
 55 60  
 Ser Glu Arg Leu Arg Gln Ala Glu Glu Ser Leu Arg Thr Val Met Phe 65  
 70 75 80  
 Leu Ser Cys Trp Gly Ser Cys 85

&lt;210&gt; 371

&lt;211&gt; 75

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..75:Ceres Seq. ID 2025474

&lt;400&gt;371

Met Lys Glu Lys Ser Val Thr Gly Val Ser Gly Gly Asp Thr Thr Leu 1  
 5 10 15  
 Arg Ser Ala Val Lys Glu Ser Val Ser Ser Pro Gln Ser Ala Ser Ser 20  
 25 30  
 Ser Ser Ser Val Arg Arg Leu Lys Gly Asp Leu Glu Ser Ser Arg Phe 35  
 40 45  
 Gly Ala Ala Ala Ser Glu Arg Leu Arg Gln Ala Glu Ser Leu Arg 50  
 55 60  
 Thr Val Met Phe Leu Ser Cys Trp Gly Ser Cys 65  
 70 75

&lt;210&gt; 372

&lt;211&gt; 2029

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..2029:Ceres Seq. ID 2025475

&lt;400&gt;372

atgggtaaga gtggtttc agctgtgaag aaagcattaa gcccagaacc aaaacaaag 60  
 aagagcaga agccacataa gtccaagaaa tggtttgta aatccaaga gctagatgtt 120  
 actaatctg gtgcagcata ttctctcgt actgtcaag acggaact aagagatt 180  
 gaggagcac agagagaca tcttactct gtgctattg caactctgc agctgcagag 240  
 gcagcgttg cagctgcca agctctgct gaagtgttc gtctctgac attatcagg 300  
 ttcccggga aatcaatgga agagatgcc gctatcaaga ttacagacgc attcagaga 360  
 tatattgcaa gaagacatt gctgctgtt agagcttgg ttaggctaaa atcttagtc 420  
 cagggaat gtgtgagac tcaagccaca tctacattgc aaagcatgca aacactgct 480  
 agagtcaat atcagattcg tgaagaagg ctccgattgt ctgaggtata acagcttta 540  
 acacgacgc tccaacaaa acacataaa gacttgata agactggaga aaattggaat 600  
 gatagtacat tgcgcggga gaagttaa gcaacatgt tgaacaagca agtgcacaa 660  
 atgagaagag aaaaacgct tgcataatga ttcatcacc agaatacatg gaaaaacta 720  
 actaaaatgg gtctcaaac attcatggac cctacaatc cgcattggg ttggagtgg 780

ctagaacgtt ggaatgctgc tgcacaaac gaaacacact cactcacacc agataatgct 840  
 gaaaagact cttctgctag gagtgtgca agcgtgcca tgtctgagt gattcacaga 900  
 gtcacaaact ttccaccag agaaagaca ccaacacgtc gaagagggtc aagcccgaga 960  
 gtgggcaag tcccaagga agactcaac agcattgtg gtttccaatc agaacacact 1020  
 tgcacatcga ggaatgacac ttgtgtatca attccatcaa ctagaagcga tgaagcttc 1080  
 acagatgatt tctctcagtc agttccagtc tacaatgccc ctacacaaagc gcgcaagca 1140  
 agagctcgat tctcaaacct tagtctctta agtcacagga agacagcga aaacggtt 1200  
 tctctcgg gatctccaa gacttaaga cggttttcag tagggattct aggccttcc 1260  
 ttcattggct agatagcag agcttttaca tctctctctt ctctctctc gatattcgtc 1320  
 agtgtgtga ttgtgggaa actttgtga gcaagagagc agaaaatgag cgaccggta 1380  
 agaaatcgc ggaatggct tcaaacgtt cgaagactat cgaatggat ggtatgcta 1440  
 aggtcttct caagatagc gccatagag agttcttaa ccttcgctgt gctttgatg 1500  
 aggttaaac acagctcag accaaatta ttgcacagta caaggaagct tatgacaga 1560  
 actatagaa ggtatgtga gttggcattg ttgacaagta caaggaagct tatgacaga 1620  
 ttgagttcc aaagtacgtt gcaagatga cctctgaata caegcnaag ttgatgctt 1680  
 tgtgtgtga actgaagaa gcaagacaga aatctctcaa ggaatctgaa cgttgtgaga 1740  
 aagaattgc tgaatccaa gagatcaga aagaactcag caccatgact gcagatgagt 1800  
 actttgaga gacccgaa ctcaaaaga agtttgatga cgaatcgt aatgacaact 1860  
 ggggatactg ataatgttc tcatctcgc ggttggaag aaaaactctt ttctcttct 1920  
 ctgtctctt actgtgatt ttgtggcga taaaacat aatagata caattcactt 1980  
 aagcagtggt gagatcttca ttccaagaa gataacgca ttgtgtttt 2029

&lt;210&gt; 373

&lt;211&gt; 449

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..449:Ceres Seq. ID 2025476

&lt;400&gt;373

Met Gly Lys Ser Trp Phe Ser Ala Val Lys Lys Lys Ala Leu Ser Pro Glu

1 5 10 15

Pro Lys Gln Lys Lys Glu Gln Lys Pro His Lys Ser Lys Lys Trp Phe

20 25 30

Gly Lys Ser Lys Lys Leu Asp Val Thr Asn Ser Gly Ala Ala Tyr Ser

35 40 45

Pro Arg Thr Val Lys Asp Ala Lys Lys Leu Lys Glu Ile Glu Gln Gln

50 55 60

Ser Arg His Ala Tyr Ser Val Ala Ile Ala Thr Ala Ala Ala Glu

65 70 75 80

Ala Ala Val Ala Ala Ala Gln Ala Ala Ala Glu Val Val Arg Leu Ser

85 90 95

Ala Leu Ser Arg Phe Pro Gly Lys Ser Met Glu Glu Ile Ala Ala Ile

100 105 110

Lys Ile Gln Thr Ala Phe Arg Gly Tyr Met Ala Arg Arg Ala Leu Arg

115 120 125

Ala Leu Arg Gly Leu Val Arg Leu Lys Ser Leu Val Gln Gly Lys Cys

130 135 140

Val Arg Arg Gln Ala Thr Ser Thr Leu Gln Ser Met Gln Thr Leu Ala

145 150 155 160

Arg Val Gln Tyr Gln Ile Arg Glu Arg Arg Leu Leu Ser Glu Asp

165 170 175

Lys Gln Ala Leu Thr Arg Gln Leu Gln Lys His Asn Lys Asp Phe

180 185 190

Asp Lys Thr Gly Glu Asn Trp Asn Asp Ser Thr Leu Ser Arg Glu Lys

195 200 205

Val Glu Ala Asn Met Leu Asn Lys Gln Val Ala Thr Met Arg Arg Glu

210 215 220

Lys Ala Leu Ala Tyr Ala Phe Ser His Gln Asn Thr Trp Lys Asn Ser 225  
 Thr Lys Met Gly Ser Gln Thr Phe Met Asp Pro Asn Asn Pro His Trp 240  
 245 250 255  
 Gly Trp Ser Trp Leu Glu Arg Trp Met Ala Ala Arg Pro Asn Glu Asn  
 260 265 270  
 His Ser Leu Thr Pro Asp Asn Ala Glu Lys Asp Ser Ser Ala Arg Ser  
 275 280 285  
 Val Ala Ser Arg Ala Met Ser Glu Met Ile Pro Arg Gly Lys Asn Leu  
 290 300  
 Ser Pro Arg Gly Lys Thr Pro Asn Ser Arg Arg Gly Ser Ser Pro Arg  
 305 310 315  
 Val Arg Gln Val Pro Ser Glu Asp Ser Asn Ser Ile Val Ser Phe Gln  
 320 325 330 335  
 Ser Glu Gln Pro Cys Asn Arg Arg His Ser Thr Cys Gly Ser Ile Pro  
 340 345 350  
 Ser Thr Arg Asp Asp Glu Ser Phe Thr Ser Ser Phe Ser Gln Ser Val  
 355 360 365  
 Pro Gly Tyr Met Ala Pro Thr Gln Ala Ala Lys Ala Arg Ala Arg Phe  
 370 375 380  
 Ser Asn Leu Ser Pro Leu Ser Ser Glu Lys Thr Ala Lys Lys Arg Leu  
 385 390 395 400  
 Ser Phe Ser Gly Ser Pro Lys Thr Val Arg Arg Phe Ser Val Gly Ile  
 405 410 415  
 Leu Gly Leu Ser Phe Ile Gly Gln Ile Arg Arg Phe Thr Ser Ser  
 420 425 430  
 Ser Ser Ser Ser Ile Phe Val Ser Val Cys Ile Leu Gly Lys Thr  
 435 440 445  
 Leu

&lt;210&gt; 374

&lt;211&gt; 344

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..344:Ceres Seq. ID 2025477

&lt;400&gt;374

Met Glu Glu Ile Ala Ala Ile Lys Ile Gln Thr Ala Phe Arg Gly Tyr

1 5 10 15

Met Ala Arg Arg Ala Leu Arg Ala Leu Arg Gly Leu Val Arg Leu Lys

20 25 30

Ser Leu Val Gln Gly Lys Cys Val Arg Arg Gln Ala Thr Ser Thr Leu

35 40 45

Gln Ser Met Gln Thr Leu Ala Arg Val Gln Tyr Gln Ile Arg Glu Arg

50 55 60

Arg Leu Arg Leu Ser Glu Asp Lys Gln Ala Leu Thr Arg Gln Leu Gln

65 70 75 80

Gln Lys His Asn Lys Asp Phe Asp Lys Thr Gly Glu Asn Trp Asn Asp

85 90 95

Ser Thr Leu Ser Arg Glu Lys Val Glu Ala Asn Met Leu Asn Lys Gln

100 105 110

Val Ala Thr Met Arg Arg Glu Lys Ala Leu Ala Tyr Ala Phe Ser His

115 120 125

Gln Asn Thr Trp Lys Asn Ser Thr Lys Met Gly Ser Gln Thr Phe Met

130 135 140

Asp Pro Asn Asn Pro His Trp Gly Trp Ser Trp Leu Glu Arg Trp Met

145 150 155 160  
 Ala Ala Arg Pro Asn Glu Asn His Ser Leu Thr Pro Asp Asn Ala Glu  
 165 170 175  
 Lys Asp Ser Ser Ala Arg Ser Val Ala Ser Arg Ala Met Ser Glu Met  
 180 185 190  
 Ile Pro Arg Gly Lys Asn Leu Ser Pro Arg Gly Lys Thr Pro Asn Ser  
 195 200 205  
 Arg Arg Gly Ser Ser Pro Arg Val Arg Gln Val Pro Ser Glu Asp Ser  
 210 215 220  
 Asn Ser Ile Val Ser Phe Gln Ser Glu Gln Pro Cys Asn Arg Arg His  
 225 230 235  
 Ser Thr Cys Gly Ser Ile Pro Ser Thr Arg Asp Asp Glu Ser Phe Thr  
 240 245 250 255  
 Ser Ser Phe Ser Gln Ser Val Pro Gly Tyr Met Ala Pro Thr Gln Ala  
 260 265 270  
 Ala Lys Ala Arg Ala Arg Phe Ser Asn Leu Ser Pro Leu Ser Ser Glu  
 275 280 285  
 Lys Thr Ala Lys Lys Arg Leu Ser Phe Ser Gly Ser Pro Lys Thr Val  
 290 295 300  
 Arg Arg Phe Ser Val Gly Ile Leu Gly Leu Ser Phe Ile Gly Gln Ile  
 305 310 315 320  
 Arg Arg Arg Phe Thr Ser Ser Ser Ser Ser Ile Phe Val Ser  
 325 330 335  
 Val Cys Ile Leu Gly Lys Thr Leu  
 340

&lt;210&gt; 375

&lt;211&gt; 328

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..328:Ceres Seq. ID 2025478

&lt;400&gt;375

Met Ala Arg Arg Ala Leu Arg Ala Leu Arg Gly Leu Val Arg Leu Lys  
 1 5 10 15  
 Ser Leu Val Gln Gly Lys Cys Val Arg Arg Gln Ala Thr Ser Thr Leu  
 20 25 30  
 Gln Ser Met Gln Thr Leu Ala Arg Val Gln Tyr Gln Ile Arg Glu Arg  
 35 40 45  
 Arg Leu Arg Leu Ser Glu Asp Lys Gln Ala Leu Thr Arg Gln Leu Gln  
 50 55 60  
 Gln Lys His Asn Lys Asp Phe Asp Lys Thr Gly Glu Asn Trp Asn Asp,  
 65 70 75 80  
 Ser Thr Leu Ser Arg Glu Lys Val Glu Ala Asn Met Leu Asn Lys Gln  
 85 90 95  
 Val Ala Thr Met Arg Arg Glu Lys Ala Leu Ala Tyr Ala Phe Ser His  
 100 105 110  
 Gln Asn Thr Trp Lys Asn Ser Thr Lys Met Gly Ser Gln Thr Phe Met  
 115 120 125  
 Asp Pro Asn Asn Pro His Trp Gly Trp Ser Trp Leu Glu Arg Trp Met  
 130 135 140  
 Ala Ala Arg Pro Asn Glu Asn His Ser Leu Thr Pro Asp Asn Ala Glu  
 145 150 155  
 Lys Asp Ser Ser Ala Arg Ser Val Ala Ser Arg Ala Met Ser Glu Met  
 160 165 170 175  
 Ile Pro Arg Gly Lys Asn Leu Ser Pro Arg Gly Lys Thr Pro Asn Ser  
 180 185 190

Arg Arg Gly Ser Ser Pro Arg Val Arg Gln Val Pro Ser Glu Asp Ser  
 195 200 205  
 Asn Ser Ile Val Ser Phe Gln Ser Glu Gln Pro Cys Asn Arg Arg His  
 210 215 220  
 Ser Thr Cys Gly Ser Ile Pro Ser Thr Arg Asp Asp Glu Ser Phe Thr  
 225 230 235 240  
 Ser Ser Phe Ser Gln Ser Val Pro Gly Tyr Met Ala Pro Thr Gln Ala  
 245 250 255  
 Ala Lys Ala Arg Ala Arg Phe Ser Asn Leu Ser Pro Leu Ser Ser Glu  
 260 265 270  
 Lys Thr Ala Lys Lys Arg Leu Ser Phe Ser Gly Ser Pro Lys Thr Val  
 275 280 285  
 Arg Arg Phe Ser Val Gly Ile Leu Gly Leu Ser Phe Ile Gly Gln Ile  
 290 295 300  
 Arg Arg Arg Phe Thr Ser Ser Ser Ser Ser Ile Phe Val Ser  
 305 310 315 320  
 Val Cys Ile Leu Gly Lys Thr Leu  
 325

&lt;210&gt; 376

&lt;211&gt; 678

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..678:Ceres Seq. ID 2025524

&lt;400&gt;376

gatttttag ggtttcaagt gaaaagagta atagcgccgc ggagccatgg ttctcaagac  
 60  
 tggctttgc cgattcagtg gccagaaaat ttacctgggt agaggatca gatttatccg  
 120  
 atggactct caggtgtttt tgtttctcaa ttcctaatgt aagggtatt tccacaacaa  
 180  
 gttagaccca tctaaagcttt gctgactgc tatgtaccga aagcagcaca agaagcgc  
 240  
 agcacaagag gctgtgaaga gaagagacg tgcactaag agccttaet caagctcgat  
 300  
 tgcgtgtct acttggggg ttattcagaa caagcagca gagaagccg agttcgtga  
 360  
 tgcgtgtga gaagtgccc tactggagt caagcagaga atcaagaaga ccaagcaga  
 420  
 gaagagcca agaaggtcgc agtatgcatc aaagcaacag aagtcacaag tgaagggaaa  
 480  
 tatccccaag agtgtgcac ccaaggtcgc taagatgggt ggtgtggag gcagagcttg  
 540  
 aatggagcta tagatagccc cactctctc tcttcaetta tctttctc ttgtttgac  
 600  
 attgtttgt ttgtcagcc atttttagt ttgtgaccag atctaataata ttcaatttat  
 660  
 gaaaactttt tgtttggc  
 678

&lt;210&gt; 377

&lt;211&gt; 164

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..164:Ceres Seq. ID 2025525

&lt;400&gt;377

Met Val Leu Lys Thr Glu Leu Cys Arg Phe Ser Gly Gln Lys Ile Tyr  
 1 5 10 15  
 Pro Gly Arg Gly Ile Arg Phe Ile Arg Ser Asp Ser Gln Val Phe Leu  
 20 25 30  
 Phe Leu Asn Ser Lys Cys Lys Arg Tyr Phe His Asn Lys Leu Lys Pro  
 35 40 45  
 Ser Lys Leu Cys Trp Thr Ala Met Tyr Arg Lys Gln His Lys Lys Asp  
 50 55 60



50  
Ala Ala Gln Glu Ala Val Lys Arg Arg Arg Arg Ala Thr Lys Lys Pro  
65 70 75 80  
Tyr Ser Arg Ser Ile Val Gly Ala Thr Leu Glu Val Ile Gln Lys Lys  
85 90 95  
Arg Ala Glu Lys Pro Glu Val Arg Asp Ala Ala Arg Glu Ala Ala Leu  
100 105 110  
Arg Glu Ile Lys Lys Glu Arg Ile Lys Lys Thr Lys Asp Glu Lys Lys Ala  
115 120 125  
Lys Lys Val Glu Tyr Ala Ser Lys Gln Gln Lys Ser Gln Val Lys Gly  
130 135 140  
Asn Ile Pro Lys Ser Ala Ala Pro Lys Ala Lys Met Gly Gly Gly  
145 150 155 160  
Gly Gly Arg Arg

<210> 378  
<211> 109  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..109;Ceres Seq. ID 2025526

<400>378  
Met Tyr Arg Lys Gln His Lys Lys Lys Asp Ala Ala Gln Glu Ala Val Lys  
1 5 10 15  
Arg Arg Arg Arg Ala Thr Lys Lys Pro Tyr Ser Arg Ser Ile Val Gly  
20 25 30  
Ala Thr Leu Glu Val Ile Gln Lys Lys Arg Ala Glu Lys Pro Glu Val  
35 40 45  
Arg Asp Ala Ala Arg Glu Ala Ala Leu Arg Glu Ile Lys Glu Arg Ile  
50 55 60  
Lys Lys Thr Lys Asp Glu Lys Lys Ala Lys Lys Val Glu Tyr Ala Ser  
65 70 75 80  
Lys Gln Gln Lys Ser Gln Val Lys Gly Asn Ile Pro Lys Ser Ala Ala  
85 90 95  
Pro Lys Ala Ala Lys Met Gly Gly Gly Gly Arg Arg  
100 105

<210> 379  
<211> 1623  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..1623;Ceres Seq. ID 2025544

<400>379  
atggcattt ctaagatgt tottctgaaa ttacatatg ttgtttcttt ctggattact  
60  
tgcgtatct tcaaatgtc aagaacgata aaacagctt ttgtggctc attcatctc  
120  
attcttgctt ccctgagtgt tgyggagcc gccattgtag gggccatcga aggtcaccc  
180  
actgaatcg gattttcca agttagtgt ctgtgagtgg ttgctgagt tatcaetgg  
240  
gtccaaact ttggaccggt gtcacatagt gatcaacct tgcctaaagt ggccttactc  
300  
aggagatgt tgaatgggaa agccattgtt ggtatggta gacetttgt ttcaaaaga  
360  
tatacatgc aaataatgc atgtgtaca agttacatgg agatttcaa ctatcagat  
420  
ttcaatcatg aaaaaaagg actatcgaag agctctattc agaacatccc gatgtctac  
480

aacccttcag aacatcaaac aaatcgaat tgctcgattt gcttacagga ttgggaagaa  
540  
ggggaagtag gaagaagct ggaagatgt ggcacacat ttcatatgaa ttgcatagat  
600  
gagtggtgc ttagaacga aacttgccc atttgaaga gacaaagaa gaagcagcag  
660  
aatcatcgc ttaaaatcca atgcatacga ctacgtctc ttctctctt tcactacct  
720  
ttctctctt ctctcgtggt ccattcgta ttctaaact ctgtcgccg cgggagagag  
780  
ttagatgact tggctggttt tcaatccat aaatactca acatcgcaa cattcgtgg  
840  
attcgtctc actcgtacc cattgtcgc tgggtgtc ggctggcaa cattcgtgg  
900  
tgttctctc ttcaagaatg cttctagatc tctctcttc gcttccctt ctgtctctg  
960  
ttctctctg aagaagaag ttgtctctc ttggaagca ccagctgctt tgggacctta  
1020  
ctctcagcc attaaagca ataattgtt ttctcttca ggtgtcttg gacttacct  
1080  
ttagactgga asgttgttt cggagacgtt cgaagatcag actgacagag tactcaaaa  
1140  
catggggag aatttgaag ctatgggtc tgaattatcc tcggtgtgga agacaacat  
1200  
cattgtgctt gatttggtt actcaagac agtgaagcag atatatgca aatactccc  
1260  
agctctctt ccagcagat cgaactatca agttcagct ttgctctaa acccaagat  
1320  
cgaagatgaa tgaattgcaa cactctagaa cactcaaatc aatctctac aggggtattt  
1380  
tggctcgaa gctgtcaaaa conataagag aaacatggt ttgtgggct taaccatct  
1440  
ttaacttact ttctatttt ctgcactacg aaaaatttg cggattatca accattaggt  
1500  
tggttacaaa tcaaaagcct ctaatacaga aaagtattt tagtctctg tccacctca  
1560  
tttgctgtg acgtatatgt attgattgta tacattattt tatttgagt aaaccattat  
1620  
gct  
1623

<210> 380  
<211> 249  
<212> PRT  
<213> Arabidopsis thaliana  
  
<220>  
<223> any n or Xaa = unknown  
  
<223> LOCATION: 1..249;Ceres Seq. ID 2025545

<400>380  
Met Ser Gly Cys Leu Asp Lys Lys Leu Ala Pro Phe Glu Glu Asp Lys  
1 5 10 15  
Arg Arg Ser Thr Arg Ile Ile Cys Leu Lys Ser Asn Arg Ile Asp Ser  
20 25 30  
Val Phe Phe Phe Phe Phe Thr Tyr Leu Ser Leu Ser Leu Ser Trp Ser  
35 40 45  
Ile Arg Tyr Ser Ile Leu Leu Val Gly Ala Gly Glu Ser Glu Met Thr  
50 55 60  
Trp Ser Val Phe Arg Ser Ile Asn Thr Pro Thr Leu Asp Leu Ser Thr  
65 70 75  
Ala Leu Arg Ser Thr Arg Thr Pro Leu Val Ala Ala Gly Val Gly Cys  
80 85 90 95  
Ala Thr Phe Ala Gly Val Ser Leu Phe Arg Met Ser Ser Arg Ser Pro  
100 105 110  
Pro Phe Ala Ser Leu Ser Val Ser Ala Ser Ser Val Lys Lys Glu Val  
115 120 125  
Val Ser Thr Glu Lys Ala Pro Ala Ala Leu Gly Pro Tyr Ser Gln Ala  
130 135 140  
Ile Lys Ala Asn Asn Leu Val Phe Leu Ser Gly Val Leu Gly Ile Ile  
145 150 155  
Pro Glu Thr Gly Lys Phe Val Ser Glu Ser Val Glu Asp Gln Thr Glu  
160 165 170 175  
Gln Val Leu Lys Asn Met Gly Glu Ile Leu Lys Ala Ser Gly Ala Asp  
180 185 190  
Tyr Ser Ser Val Val Lys Thr Thr Ile Met Leu Ala Asp Leu Ala Asp  
195 200 205  
Phe Lys Thr Val Asn Glu Ile Tyr Ala Lys Tyr Phe Pro Ala Pro Ser  
210 215 220  
Pro Ala Arg Ser Thr Tyr Gln Val Ala Ala Leu Pro Leu Asn Ala Lys

225 Ile Glu Ile Glu Cys Ile Ala Thr Leu 230 235 240  
245

<210> 381

<211> 211

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..211:Ceres Seq. ID 2025546

<400>381

Met Gly Ile Ser Lys Met Ile Leu Leu Lys Phe Thr Tyr Val Val Ser

1

Phe Trp Ile Thr Cys Val Ile Phe Lys Met Ser Arg Thr Ile Lys Arg

20

Ala Leu Ileu Gly Ser Phe Ile Phe Ile Leu Ala Ser Ala Ser Val Val

35

Val Ala Ala Ile Val Gly Ala Ile Glu Gly His Thr Thr Asp Ile Gly

50

Phe Leu Gln Gly Ser Val Leu Gly Val Val Ala Gly Val Ile Thr Ala

65

Val Gln Leu Phe Gly Pro Val Leu His Ser Asp Gln Pro Leu Ser Lys

85

Val Ala Leu Leu Arg Arg Val Val Asn Gly Lys Ala Ile Met Gly Leu

100

Val Arg Pro Phe Val Leu Lys Ala Tyr Gln Trp Gln Ile Ile Ala Leu

110

Asp Thr Ser Tyr Met Glu Ser Ser Asn Leu Tyr Asp Phe Asn His Glu

130

Lys Lys Gly Leu Ser Lys Ser Ile Gln Asn Ile Pro Met Phe Tyr

145

Asn Arg Ser Glu His Gln Thr Lys Ser Ser Cys Ser Ile Cys Leu Gln

165

Asp Trp Glu Glu Gly Glu Val Gly Arg Lys Leu Ala Arg Cys Gly His

180

Thr Phe His Met Asn Cys Ile Asp Glu Trp Leu Leu Arg Gln Glu Thr

195

Cys Pro Ile

210

<210> 382

<211> 206

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..206:Ceres Seq. ID 2025547

<400>382

Met Ile Leu Leu Lys Phe Thr Tyr Val Val Ser Phe Trp Ile Thr Cys

1

Val Ile Phe Lys Met Ser Arg Thr Ile Lys Arg Ala Leu Leu Gly Ser

20

Phe Ile Phe Ile Leu Ala Ser Ala Ser Val Val Val Ala Ala Ile Val

35

Gly Ala Ile Glu Gly His Thr Thr Asp Ile Gly Phe Leu Gln Gly Ser

50

Val Leu Gly Val Val Ala Gly Val Ile Thr Ala Val Gln Leu Phe Gly

65

Pro Val Leu His Ser Asp Gln Pro Leu Ser Lys Val Ala Leu Leu Arg

85

Arg Val Val Asn Gly Lys Ala Ile Met Gly Leu Val Arg Pro Phe Val

100

Leu Lys Ala Tyr Gln Trp Gln Ile Ile Ala Leu Asp Thr Ser Tyr Met

115

Glu Ser Ser Asn Leu Tyr Asp Phe Asn His Glu Lys Lys Gly Leu Ser

130

Lys Ser Ser Ile Gln Asn Ile Pro Met Phe Tyr Asn Arg Ser Glu His

145

Gln Thr Lys Ser Ser Cys Ser Ile Cys Leu Gln Asp Trp Glu Glu Gly

165

Glu Val Gly Arg Lys Leu Ala Arg Cys Gly His Thr Phe His Met Asn

180

Cys Ile Asp Glu Trp Leu Leu Arg Gln Glu Thr Cys Pro Ile

195

<210> 383

<211> 542

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..542:Ceres Seq. ID 2026207

<400>383

tcccaaat ctctctctc ttctccattg ttctctcta atggaacca ccgagaatc

60

taagtacta gaactacta gtaccagaa gatgagacc aaatcacga atagttaac

120

agttatggc gagaagatgg atgttgaga gttgtatca gagttatga aagtttcag

180

tttgcctgg gatcccgaga gacactcat cacagctgag tcttaagac gaactcagg

240

gactctgga attgaagta ttagcaaga agatctcaa ggaatgta gagaagaga

300

ccctgagatg atggaagcg ccgaacttg gtggagaaa gaactcacc aagaactatg

360

taatcaaat ctctctcta tgcctgaat gatcgtctc ctctcttg tattctctt

420

ttactctac aaacactata aatgtttctg aaataaaat gaagtaatga ttcttgatc

480

cc

<210> 384

<211> 148

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..148:Ceres Seq. ID 2026208

<400>384

His Gln Asn Leu Phe Ser Leu Phe His Cys Leu Pro Leu Met Glu Pro

1

Thr Glu Lys Ser Met Leu Leu Glu Thr Thr Ser Thr Thr Lys Met Glu

15

Thr Lys Tyr Glu Asp Met Leu Pro Val Met Ala Glu Lys Met Asp Val

20

Glu Glu Phe Val Ser Glu Leu Cys Lys Gly Phe Ser Leu Leu Ala Asp

25

45

Pro Glu Arg His Leu Ile Thr Ala Glu Ser Leu Arg Arg Ser Gly 60  
 65 Ile Leu Gly Ile Glu Gly Met Ser Lys Glu Asp Ala Gln Gly Met Val 75  
 85 Arg Glu Gly Asp Leu Asp Gly Ala Leu Asn Gln Thr Glu Phe 95  
 100 Cys Val Leu Met Val Arg Leu Ser Pro Glu Met Met Glu Asp Ala Glu 110  
 115 Thr Trp Leu Glu Lys Ala Leu Thr Gln Glu Leu Cys Asn His Asn Leu 125  
 130 Ser Ser Met Pro 135  
 145

<210> 385  
 <211> 135  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..135:Ceres Seq. ID 2026209

Met Glu Pro Thr Glu Lys Ser Met Leu Leu Glu Thr Thr Ser Thr Thr 15  
 1 Lys Met Glu Thr Lys Tyr Glu Asp Met Leu Pro Val Met Ala Glu Lys 30  
 20 Met Asp Val Glu Glu Phe Val Ser Glu Leu Cys Lys Gly Phe Ser Leu 45  
 35 Leu Ala Asp Pro Glu Arg His Leu Ile Thr Ala Glu Ser Leu Arg Arg 60  
 50 Asn Ser Gly Ile Leu Gly Ile Glu Gly Met Ser Lys Glu Asp Ala Gln 75  
 65 Gly Met Val Arg Glu Gly Asp Leu Asp Gly Asp Gly Ala Leu Asn Gln 90  
 85 Thr Glu Phe Cys Val Leu Met Val Arg Leu Ser Pro Glu Met Met Glu 110  
 100 Asp Ala Glu Thr Trp Leu Glu Lys Ala Leu Thr Gln Glu Leu Cys Asn 125  
 115 His Asn Leu Ser Ser Met Pro 130  
 135

<210> 386  
 <211> 128  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..128:Ceres Seq. ID 2026210

Met Leu Leu Glu Thr Thr Ser Thr Thr Lys Met Glu Thr Lys Tyr Glu 15  
 1 Asp Met Leu Pro Val Met Ala Glu Lys Met Asp Val Glu Glu Phe Val 30  
 20 Ser Glu Leu Cys Lys Gly Phe Ser Leu Leu Ala Asp Pro Glu Arg His 45  
 35

Leu Ile Thr Ala Glu Ser Leu Arg Arg Asn Ser Gly Ile Leu Gly Ile 60  
 50 Glu Gly Met Ser Lys Glu Asp Ala Gln Gly Met Val Arg Glu Gly Asp 75  
 65 Leu Asp Gly Asp Gly Ala Leu Asn Gln Thr Glu Phe Cys Val Leu Met 80  
 85 Val Arg Leu Ser Pro Glu Met Met Glu Asp Ala Glu Thr Trp Leu Glu 95  
 100 Lys Ala Leu Thr Gln Glu Leu Cys Asn His Asn Leu Ser Ser Met Pro 110  
 115 120 125

<210> 387  
 <211> 352  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..552:Ceres Seq. ID 2026982

agaagatata aaacgctttt tcttaaatga tggatgtctt cttttatat agtgcacat 60  
 tctctatag cactctacag tctgcacact taccacttct ttcgaaatct cttctatcc 120  
 tctcaaaa tgaagctctc tgtcgctttt atctcgctg cttctctctt gttctatgta 180  
 ttcattacca caggatggg tccagtcacc gtgagggcac gcactgtga gtcacaaagc 240  
 catagttcca agggctccat tctgagcaca cacacactgt caaacgtgtg ccaacagaa 300  
 gcttcgctgc gagtaaatg ccgtggtatc cgtcgctgtt gctaactgac aagacactgc 360  
 tcatcaatcc attctcatga ctcaaatctt cgatccatcg tcaagtgttt actctttct 420  
 tctcaaatc ttcgctacgg taccatgtcg taccgtacat gagtgttttc tgaataagt 480  
 attgtgttg tgtgttccg gttttaatgt aatgtaaat caattaatgg ctttaatat 540  
 attgtattat gg 552

<210> 388  
 <211> 120  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..120:Ceres Seq. ID 2026983

Met Lys Tyr Lys Asn Val Phe Leu Lys Leu Trp Met Ser Ser Leu Tyr 15  
 1 Ile Val Pro Thr Ser Ser Leu Ala Phe Tyr Ser Leu His Thr Tyr His 30  
 20 Phe Phe Arg Asn Leu Ser Leu Ser Lys Met Lys Leu Ser Val 45  
 35 Arg Phe Ile Ser Ala Ala Leu Leu Phe Met Val Phe Ile Ala Thr 60  
 50 Gly Met Gly Pro Val Thr Val Glu Ala Arg Thr Cys Glu Ser Lys Ser 75  
 65 His Arg Phe Lys Gly Pro Cys Val Ser Thr His Asn Cys Ala Asn Val 80  
 85 Cys His Asn Glu Gly Phe Gly Gly Lys Cys Arg Gly Phe Arg Arg 110  
 100 Arg Cys Tyr Cys Thr Arg His Cys 115  
 120

<210> 389  
 <211> 109  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..109:Ceres Seq. ID 2026984

<400>389  
 Met Ser Ser Leu Tyr Ile Val Pro Thr Ser Ser Leu Ala Phe Tyr Ser  
 1 5 10 15  
 Leu His Thr Tyr His Phe Phe Arg Asn Leu Ser Leu Ser Leu Ser Lys  
 20 25 30  
 Met Lys Leu Ser Val Arg Phe Ile Ser Ala Ala Leu Leu Leu Phe Met  
 35 40 45  
 Val Phe Ile Ala Thr Gly Met Gly Pro Val Thr Val Glu Ala Arg Thr  
 50 55 60  
 Cys Glu Ser Lys Ser His Arg Phe Lys Gly Pro Cys Val Ser Thr His  
 65 70 75 80  
 Asn Cys Ala Asn Val Cys His Asn Glu Gly Phe Gly Gly Lys Cys  
 85 90 95  
 Arg Gly Phe Arg Arg Arg Cys Tyr Cys Thr Arg His Cys  
 100 105

<210> 390  
 <211> 77  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..77:Ceres Seq. ID 2026985

<400>390  
 Met Lys Leu Ser Val Arg Phe Ile Ser Ala Ala Leu Leu Leu Phe Met  
 1 5 10 15  
 Val Phe Ile Ala Thr Gly Met Gly Pro Val Thr Val Glu Ala Arg Thr  
 20 25 30  
 Cys Glu Ser Lys Ser His Arg Phe Lys Gly Pro Cys Val Ser Thr His  
 35 40 45  
 Asn Cys Ala Asn Val Cys His Asn Glu Gly Phe Gly Gly Lys Cys  
 50 55 60  
 Arg Gly Phe Arg Arg Arg Cys Tyr Cys Thr Arg His Cys  
 65 70 75

<210> 391  
 <211> 66  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..66:Ceres Seq. ID 2026986

<400>391  
 Met Asp Val Phe Ser Leu Tyr Ser Ala His Phe Leu Ile Ser Ile Leu  
 1 5 10 15

Gln Ser Ala His Leu Ser Leu Ser Lys Ser Leu Ser Ile Pro Leu  
 20 25 30  
 Lys Asn Glu Ala Leu Cys Ala Phe Tyr Leu Arg Cys Ser Ser Leu Val  
 35 40 45  
 His Gly Ile His Cys His Arg Asp Gly Ser Ser His Arg Gly Gly Thr  
 50 55 60  
 His Val  
 65

<210> 392  
 <211> 1760  
 <212> DNA  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..1760:Ceres Seq. ID 2027300

<400>392  
 ctggttcat gagagaaca aaaaatttta tcttttgac aacgtgtgtg gtgtgtgatg 60  
 ttgtcagca tggcaagac agtacaacy atctctctct ctittcttct tctctgtgt 120  
 ctgtcttaa atagatctc ttggaacga agcagcttg tctgtttctg ataccatgg 180  
 ctttacctt gtgtctaaag actgagctt taagtctcc gccacttga ttaattttt 240  
 ccgagctgt tatgaagtc gaacagctt tgcgcttcc jacaagctt cagcttactc 300  
 gaaaagctg tcgagagaga ttgagagctc agctatttt cagtttccct ctgctgttc 360  
 tgacaagaaa cgtctgagct gagaacaga aacagctcaa acaagaactt ctgagacca 420  
 ttgagctct tgaactgtgt gctacgctt cgcctgatga ccagcttcgg attgactagt 480  
 tagcgcgtaa agtggaagca gtttaaccga ccaaggagcc ttgaaagctt gatttgta 540  
 atggaatag ggagctcatt tatacaact ctgtctgat ttgcaggca aagaaccaa 600  
 ggttttaag atcataacc aactaccaat ctatactgt ggatacctt aagtgcaaa 660  
 acatggagac ttggcttct tataactcgg taactggaga cataaaacc ctcaatcga 720  
 agaaggtgc tgtgaactc caaggttta aaattctcgg atttattcct ataaaagcac 780  
 ctgatagcgc cgcgtgtgaa ctgtgatta cctatgtgga cgagaacta cgtttatcaa 840  
 gagatcttc attcatgtct catgtctggt tccatgacct ttctacattg ttctgttga 900  
 ttatgaatt tcaataaca ctacaagta aaccaattag aagtgagttg gttatgttc 960  
 ggtacaaaga tcacaacggg aqaagattg asagattgtc aaagcgatgg ctgagttgg 1020  
 ttgttctcgt cgaacacaga aqaagagct tcacgaaac gaccacgca ctcttaaaa 1080  
 ctccatgga aaagcttctg cctgttttga tatcatggc ttaccttgg tgcctaaaga 1140  
 ctggagttt aacgtctcgg gccgcggat ttaactatcc atcgacaggt ggcttgcgg 1200  
 ttccagaaa gttactagt attcgaaaag gtgacggga aagattgaga atccagctg 1260  
 ttctcagett tctccaaga aacgtggag ccgagagcg gaaacagctc aaacatgac 1320  
 ttctcagc cattagct ctgaaactg gtgcaactgc ctccctgat gatcagctc 1380  
 tgattgatca gttagcagct aaggtggaag cagttaaac acccaaggag cctctgaat 1440  
 ctgattgat caatgggaaa tggagctca ttacacac acctcgtcgg atttgaag 1500  
 caaagaacc aagttctta agatcgttaa ctaactacca atgtatcaat atgtatcac 1560  
 taaagtgca aagaatggag acttgccct tctataactc ggttaactgga gacttgacac 1620  
 cctcaactc gaagcaggtt gctgtgaac ttcaaggtt taaattctc ggtttatic 1680  
 cggtaaaagc acctgatggt actgacggg gtgaactaga gattacctat gtgagcaggg 1740  
 aactacgga caaacttga 1760

<210> 393  
 <211> 527  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..527:Ceres Seq. ID 2027301

<400>393  
Met Ala Leu Pro Ser Cys Leu Lys Thr Gly Ala Leu Met Ser Pro Ala  
1 5 10 15  
Thr Gly Phe Asn Phe Ser Gly Ser Leu Met Lys Ser Asp Ser Gly Phe  
20 25 30  
Ala Val Pro Thr Lys Leu Gln Ser Thr Arg Lys Gly Asp Arg Glu Arg  
35 40 45  
Leu Arg Val Gln Ala Ile Phe Ser Phe Pro Pro Ala Phe Leu Thr Arg  
50 55 60  
Asn Gly Arg Ala Glu Lys Gln Lys Gln Leu Lys Gln Leu Leu Glu  
65 70 75 80  
Ala Ile Glu Pro Leu Glu Arg Gly Ala Thr Ala Ser Pro Asp Asp Gln  
85 90 95  
Leu Arg Ile Asp Gln Leu Ala Arg Lys Val Glu Ala Val Asn Pro Thr  
100 105 110 115  
Lys Glu Pro Leu Lys Ser Asp Leu Val Asn Gly Lys Trp Glu Leu Ile  
120 125  
Tyr Thr Thr Ser Ala Ser Ile Leu Gln Ala Lys Lys Pro Arg Phe Leu  
130 135 140  
Arg Ser Ile Thr Asn Tyr Gln Ser Ile Asn Val Asp Thr Leu Lys Val  
145 150 155 160  
Gln Asn Met Glu Thr Trp Phe Tyr Asn Ser Val Thr Gly Asp Ile  
165 170 175  
Lys Pro Leu Asn Ser Lys Lys Val Ala Val Lys Leu Gln Val Phe Lys  
180 185 190  
Ile Leu Gly Phe Ile Pro Ile Lys Ala Pro Asp Ser Ala Arg Gly Glu  
195 200 205  
Leu Glu Ile Thr Tyr Val Asp Glu Glu Leu Arg Leu Ser Arg Asp Leu  
210 215 220  
Ser Phe Met Ser His Val Trp Phe His Asp Leu Ser Thr Leu Phe Leu  
225 230 235 240  
Leu Ile Tyr Glu Ser Ser Ile Thr Leu Gln Val Lys Pro Ile Arg Ser  
245 250 255  
Glu Leu Val Met Val Arg Tyr Lys Asp His Asn Arg Arg Arg Ile Glu  
260 265 270  
Glu Trp Ser Lys Arg Trp Ser Glu Leu Val Gly Ser Val Glu Thr Gly  
275 280 285  
Arg Lys Ser Leu Ser Glu Thr Thr Gln Pro Leu Phe Lys Thr Pro Leu  
290 295 300  
Glu Lys Leu Cys Pro Val Leu Ile Ser Met Ala Leu Pro Trp Cys Leu  
305 310 315 320  
Lys Thr Gly Val Leu Thr Ser Pro Ala Ala Gly Phe Asn His Pro Ser  
325 330 335  
Asp Ser Gly Phe Ala Val Pro Thr Lys Leu Leu Ser Ile Arg Lys Gly  
340 345 350  
Asp Arg Glu Arg Leu Arg Ile Gln Ala Val Phe Ser Phe Pro Pro Arg  
355 360 365  
Asn Gly Gly Ala Glu Lys Arg Lys Gln Leu Lys His Glu Leu Val Glu  
370 375 380  
Ala Ile Glu Pro Leu Glu Arg Gly Ala Thr Ala Ser Pro Asp Asp Gln  
385 390 395 400  
Leu Leu Ile Asp Gln Leu Ala Arg Lys Val Glu Ala Val Asn Pro Thr  
405 410 415  
Lys Glu Pro Leu Lys Ser Asp Leu Ile Asn Gly Lys Trp Glu Leu Ile  
420 425 430  
Tyr Thr Thr Ser Ala Ala Ile Leu Gln Ala Lys Lys Pro Arg Phe Leu  
435 440 445  
Arg Ser Leu Thr Asn Tyr Gln Cys Ile Asn Met Asp Thr Leu Lys Val  
450 455 460  
Gln Arg Met Glu Thr Trp Pro Phe Tyr Asn Ser Val Thr Gly Asp Leu  
465 470 475 480

Thr Pro Leu Asn Ser Lys Thr Val Ala Val Lys Leu Gln Val Phe Lys  
485 490 495  
Ile Leu Gly Phe Ile Pro Val Lys Ala Pro Asp Gly Thr Ala Arg Gly  
500 505 510  
Glu Leu Glu Ile Thr Tyr Val Asp Glu Glu Leu Arg Asp Lys Leu  
515 520 525  
<210> 394  
<211> 515  
<212> PRT  
<213> Arabidopsis thaliana  
<220>  
<223> any n or Xaa = unknown  
<223> LOCATION: 1..515:Ceres Seq. ID 2027302  
<400>394  
Met Ser Pro Ala Thr Gly Phe Asn Phe Ser Gly Ser Leu Met Lys Ser  
1 5 10 15  
Asp Ser Gly Phe Ala Val Pro Thr Lys Leu Gln Ser Thr Arg Lys Gly  
20 25 30  
Asp Arg Glu Arg Leu Arg Val Gln Ala Ile Phe Ser Phe Pro Pro Ala  
35 40 45  
Phe Leu Thr Arg Asn Gly Arg Ala Glu Lys Gln Lys Gln Leu Lys Gln  
50 55 60  
Glu Leu Leu Glu Ala Ile Glu Pro Leu Glu Arg Gly Ala Thr Ala Ser  
65 70 75 80  
Pro Asp Asp Gln Leu Arg Ile Asp Gln Leu Ala Arg Lys Val Glu Ala  
85 90 95  
Val Asn Pro Thr Lys Glu Pro Leu Lys Ser Asp Leu Val Asn Gly Lys  
100 105 110  
Trp Glu Leu Ile Tyr Thr Thr Ser Ala Ser Ile Leu Gln Ala Lys Lys  
115 120 125  
Pro Arg Phe Leu Arg Ser Ile Thr Asn Tyr Gln Ser Ile Asn Val Asp  
130 135 140  
Thr Leu Lys Val Gln Asn Met Glu Thr Trp Phe Tyr Asn Ser Val  
145 150 155 160  
Thr Gly Asp Ile Lys Pro Leu Asn Ser Lys Lys Val Ala Val Lys Leu  
165 170 175  
Gln Val Phe Lys Ile Leu Gly Phe Ile Pro Ile Lys Ala Pro Asp Ser  
180 185 190  
Ala Arg Gly Glu Leu Glu Ile Thr Tyr Val Asp Glu Leu Arg Leu  
195 200 205  
Ser Arg Asp Leu Ser Phe Met Ser His Val Trp Phe His Asp Leu Ser  
210 215 220  
Thr Leu Phe Leu Leu Ile Tyr Glu Ser Ser Ile Thr Leu Gln Val Lys  
225 230 235 240  
Pro Ile Arg Ser Glu Leu Val Met Val Arg Tyr Lys Asp His Asn Arg  
245 250 255  
Arg Arg Ile Glu Glu Trp Ser Lys Arg Trp Ser Glu Leu Val Gly Ser  
260 265 270  
Val Glu Thr Gly Arg Lys Ser Leu Ser Glu Thr Thr Gln Pro Leu Phe  
275 280 285  
Lys Thr Pro Leu Glu Lys Leu Cys Pro Val Leu Ile Ser Met Ala Leu  
290 295 300  
Pro Trp Cys Leu Lys Thr Gly Val Leu Thr Ser Pro Ala Ala Gly Phe  
305 310 315 320  
Asn His Pro Ser Asp Ser Gly Phe Ala Val Pro Thr Lys Leu Ser  
325 330 335  
Ile Arg Lys Gly Asp Arg Glu Arg Leu Arg Ile Gln Ala Val Phe Ser

340 Phe Pro Pro Arg Asn Gly Gly Ala Glu Lys Arg Lys Gln Leu Lys His 350  
 355 Glu Leu Val Glu Ala Ile Glu Pro Leu Glu Arg Gly Ala Thr Ala Ser 365  
 370 Pro Asp Asp Gln Leu Leu Ile Asp Gln Leu Ala Arg Lys Val Glu Ala 380  
 385 Val Asn Pro Thr Lys Glu Pro Leu Lys Ser Asp Leu Ile Asn Gly Lys 400  
 405 Trp Glu Leu Ile Tyr Thr Thr Ser Ala Ala Ile Leu Gln Ala Lys Lys 415  
 420 Pro Arg Phe Leu Arg Ser Leu Thr Asn Tyr Gln Cys Ile Asn Met Asp 430  
 435 Thr Leu Lys Val Gln Arg Met Glu Thr Trp Pro Phe Tyr Asn Ser Val 445  
 450 Thr Gly Asp Leu Thr Pro Leu Asn Ser Lys Thr Val Ala Val Lys Leu 460  
 465 Gln Val Phe Lys Ile Leu Gly Phe Ile Pro Val Lys Ala Pro Asp Gly 475  
 485 Thr Ala Arg Gly Glu Leu Glu Ile Thr Tyr Val Asp Glu Leu Arg 495  
 500 Asp Lys Leu 510

<210> 395  
 <211> 502  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..502:Ceres Seq. ID 2027303

<400>395  
 1 Met Lys Ser Asp Ser Gly Phe Ala Val Pro Thr Lys Leu Gln Ser Thr 15  
 5 Arg Lys Gly Asp Arg Glu Arg Leu Arg Val Gln Ala Ile Phe Ser Phe 30  
 20 Pro Pro Ala Phe Leu Thr Arg Asn Gly Arg Ala Glu Lys Gln Lys Gln 45  
 35 Leu Lys Gln Glu Leu Leu Glu Ala Ile Glu Pro Leu Glu Arg Gly Ala 60  
 50 Thr Ala Ser Pro Asp Asp Gln Leu Arg Ile Asp Gln Leu Ala Arg Lys 75  
 65 Val Glu Ala Val Asn Pro Thr Lys Glu Pro Leu Lys Ser Asp Leu Val 90  
 85 Asn Gly Lys Trp Glu Leu Ile Tyr Thr Ser Ala Ser Ile Leu Gln 110  
 100 Ala Lys Lys Pro Arg Phe Leu Arg Ser Ile Thr Asn Tyr Gln Ser Ile 125  
 115 Asn Val Asp Thr Leu Lys Val Gln Asn Met Glu Thr Trp Pro Phe Tyr 140  
 130 Asn Ser Val Thr Gly Asp Ile Lys Pro Leu Asn Ser Lys Lys Val Ala 155  
 145 Val Lys Leu Gln Val Phe Lys Ile Leu Gly Phe Ile Pro Ile Lys Ala 170  
 165 Pro Asp Ser Ala Arg Gly Glu Leu Glu Ile Thr Tyr Val Asp Glu Gln 185  
 180 Leu Arg Leu Ser Arg Asp Leu Ser Phe Met Ser His Val Trp Phe His 200  
 195

Asp Leu Ser Thr Leu Phe Leu Leu Ile Tyr Glu Ser Ser Ile Thr Leu 210  
 215 Gln Val Lys Pro Ile Arg Ser Glu Leu Val Met Val Arg Tyr Lys Asp 220  
 225 His Asn Arg Arg Arg Ile Glu Glu Trp Ser Lys Arg Trp Ser Glu Leu 240  
 245 Val Gly Ser Val Glu Thr Gly Arg Lys Ser Leu Ser Glu Thr Thr Gln 255  
 260 Pro Leu Phe Lys Thr Pro Leu Glu Lys Leu Cys Pro Val Leu Ile Ser 270  
 275 Met Ala Leu Pro Trp Cys Leu Lys Thr Gly Val Leu Thr Ser Pro Ala 285  
 290 Ala Gly Phe Asn His Pro Ser Asp Ser Gly Phe Ala Val Pro Thr Lys 300  
 305 Leu Leu Ser Ile Arg Lys Gly Asp Arg Glu Arg Leu Arg Ile Gln Ala 315  
 320 Val Phe Ser Phe Pro Pro Arg Asn Gly Gly Ala Glu Lys Arg Lys Gln 330  
 335 Leu Lys His Glu Leu Val Glu Ala Ile Glu Pro Leu Glu Arg Gly Ala 350  
 355 Thr Ala Ser Pro Asp Asp Gln Leu Leu Ile Asp Gln Leu Ala Arg Lys 365  
 370 Val Glu Ala Val Asn Pro Thr Lys Glu Pro Leu Lys Ser Asp Leu Ile 380  
 385 Asn Gly Lys Trp Glu Leu Ile Tyr Thr Thr Ser Ala Ala Ile Leu Gln 395  
 400 Ala Lys Lys Pro Arg Phe Leu Arg Ser Leu Thr Asn Tyr Gln Cys Ile 410  
 415 Asn Met Asp Thr Leu Lys Val Gln Arg Met Glu Thr Trp Pro Phe Tyr 430  
 435 Asn Ser Val Thr Gly Asp Leu Thr Pro Leu Asn Ser Lys Thr Val Ala 445  
 450 Val Lys Leu Gln Val Phe Lys Ile Leu Gly Phe Ile Pro Val Lys Ala 460  
 465 Pro Asp Gly Thr Ala Arg Gly Glu Leu Glu Ile Thr Tyr Val Asp Glu 475  
 485 Glu Leu Arg Asp Lys Leu 490

<210> 396

<211> 539

<212> DNR

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..539:Ceres Seq. ID 2027375

<400>396  
 actgtgttcg ataataatgt cgacaggaga agcgatacca agagtcgccc ttgtcgtttt 60  
 catctcaac ggaactcea tcttattagg tgcgcgcgt tctcaatcg gcaactcac 120  
 ttgcgtctt ccggtggcc acctgaatt cggagagagc ttgaagaat gtgcagcgag 180  
 agaaatattc gaggaacacag gtctaaagat tgaagaatg aagcttttga ctgttaccaa 240  
 caatgcttc aaagaacac caactgcac acactacgtc ttgttttga taactgtcgtt 300  
 gtgtgtgat ccaagtcaag accgaagaa tatggaacca gagaagtgtg aagatggga 360  
 ttgttatgat tgggaagatc taccaaagc ttgttttgg ccaattgga aattgtttgg 420  
 aagtggttc atccttcca ctatgttgg tggagactaa tagatgaag agttaatgat 480  
 tgatttggga ttgaatgttg cacaatttgg gcatttggc tagtgttatg atctygttt 539

<210> 397

<211> 152  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..152:Ceres Seq. ID 2027376

<400>397  
 1 Leu Cys Ser Ile Ile Met Ser Thr Gly Glu Ala Ile Pro Arg Val Ala 15  
 5 Val Val Val Phe Ile Leu Asn Gly Asn Ser Ile Leu Leu Gly Arg Arg 15  
 20 Arg Ser Ser Ile Gly Asn Ser Thr Phe Ala Leu Pro Gly Gly His Leu 30  
 35 Glu Phe Gly Glu Ser Phe Glu Glu Cys Ala Ala Arg Glu Val Met Glu 45  
 50 Glu Thr Gly Leu Lys Ile Glu Lys Met Lys Leu Thr Val Thr Asn 60  
 65 Asn Val Phe Lys Glu Ala Pro Thr Pro Ser His Tyr Val Ser Val Ser 75  
 85 Ile Arg Ala Val Leu Val Asp Pro Ser Gln Glu Pro Lys Asn Met Glu 95  
 100 Pro Glu Lys Cys Glu Gly Trp Asp Trp Tyr Asp Trp Glu Asn Leu Pro 110  
 115 Lys Pro Leu Phe Thr Pro Leu Glu Lys Leu Phe Gly Ser Gly Phe Asn 125  
 130 Pro Phe Thr His Gly Gly Asp 140  
 145

<210> 398  
 <211> 147  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..147:Ceres Seq. ID 2027377

<400>398  
 1 Met Ser Thr Gly Glu Ala Ile Pro Arg Val Ala Val Val Phe Ile 15  
 5 Leu Asn Gly Asn Ser Ile Leu Leu Gly Arg Arg Ser Ser Ile Gly 15  
 20 Asn Ser Thr Phe Ala Leu Pro Gly Gly His Leu Glu Phe Gly Glu Ser 30  
 35 Phe Glu Glu Cys Ala Ala Arg Glu Val Met Glu Glu Thr Gly Leu Lys 45  
 50 Ile Glu Lys Met Lys Leu Thr Val Thr Asn Asn Val Phe Lys Glu 60  
 65 Ala Pro Thr Pro Ser His Tyr Val Ser Val Ser Ile Arg Ala Val Leu 75  
 85 Val Asp Pro Ser Gln Glu Pro Lys Asn Met Glu Pro Glu Lys Cys Glu 95  
 100 Gly Trp Asp Trp Tyr Asp Trp Glu Asn Leu Pro Lys Pro Leu Phe Trp 110  
 115 Pro Leu Glu Lys Leu Phe Gly Ser Gly Phe Asn Pro Phe Thr His Gly 125  
 130 Gly Gly Asp 140

145

<210> 399  
 <211> 90  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..90:Ceres Seq. ID 2027378

<400>399  
 1 Met Glu Glu Thr Gly Leu Lys Ile Glu Lys Met Lys Leu Leu Thr Val 15  
 5 Thr Asn Asn Val Phe Lys Glu Ala Pro Thr Pro Ser His Tyr Val Ser 15  
 20 Val Ser Ile Arg Ala Val Leu Val Asp Pro Ser Gln Glu Pro Lys Asn 30  
 35 Met Glu Pro Glu Lys Cys Glu Gly Trp Asp Trp Tyr Asp Trp Glu Asn 45  
 50 Leu Pro Lys Pro Leu Phe Thr Pro Leu Glu Lys Leu Phe Gly Ser Gly 60  
 65 Phe Asn Pro Phe Thr His Gly Gly Asp 75  
 85

<210> 400  
 <211> 3983  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..3983:Ceres Seq. ID 2028729

<400>400  
 atgaacacga cgcggtttca ctccgatact ccgcctoga ggatccagcg taagtctgtt 60  
 gtccaggttg ttgaagctcg taatatctc cctaagatg gtcaaggag ctctagcgtt 120  
 taagtctgtg tgcatttcga tgcctgaag aaacgaact ccactaagt ccgtgacctt 180  
 aacctattt ggaacgagat gcttgatttc gcgtctccg atcccaaaa catggtattac 240  
 gacggctcg acatcgaggt ttacaacat aaaaagatttg gtaacggagg tggccggaga 300  
 aatcttttc tggtaggtt taagatctat ggaagccagt tctcgcga aggtgaagaa 360  
 ggtcttggt atttccctt ggagaaga agtgtgttca gctggattcg cgcgagatt 420  
 ggaactaaa ttactatta cgaagaagc gccagaga acacggcggt tggaggtgga 480  
 gacggcaac acaacacga acagcaaca ttctccgc cgcaacaga agccatgaa 540  
 caacacacc agcaacaatt tcatctccg ccgcagcaga tgatgatat accacggag 600  
 aaacctaatg tagtttggt tgaagaagt aggttttcc aatcggtcca gactcagcg 660  
 tatcacaga cactacaga accctccggt gtattgttg aagaatccc accgacat 720  
 gtaatgcaag gtccaaatga taacctctt caccgaatg ataacctcc tcaacggcca 780  
 ccgtctccg gccacctcc atcggctggg gaatacat attatccac ggaagtggg 840  
 agatgcaag taggaagacc tcccgaggga gatagaatta ggtttacga gagacacgg 900  
 gagaagaaga tctcatctc ttacaattt gttgagccaa tgcagtatct ctctgttgg 960  
 attgtgaag cgcgtgctt accacatac gagaagcgt atgttaagt acgagctcg 1020  
 accatttcc taggttcaa accggcgtt accggcccg ggaatcgtt tgaatcccg 1080  
 gagtgaatc aggttttgc tcttggtcat aaccggtcg attcctgt aactgttgcg 1140  
 actctgaga tctctgttg ggaatcgcc tccgagatt ttctcgagg agtttgtttt 1200  
 gatctctg agtttcggt tctgacccg ccgatagtc cgttctccc tcaagtgtat 1260  
 cgcctcgaag gctccggcg ggaagaac tctggaga ttcccgga catcagctc 1320  
 tctgtttgga ttggtactca ggtatagag gcatttcgg aggtttggag ctctgatgt 1380  
 1440

ccgcagtag ctacacacgc ttctaaggtg tatcaatcgc cgaactttg gtacttgaga  
gtgacgttc ttgagccaca gaattacac atagctocta atctccgcgc gttgactgcg  
cttgagattc gttggaagc gaattatggg tttaagtcgc ccgtatacag aagagctca  
atgaataacc acagtggttc gttctatgg catgagata tgatctttg tctgtagaga  
ccgttggaag attgcttggt tctgaigtg gaagccgga gaataaaga agcaacatt  
ctagacatg ccatgatacc agtagctcc atcagcagc gaattatga cgcgtttg  
ccgtcgaat ggcacactt ggaagaga gttgaggtg gaggtagg agggagact  
ggagtggtg gttgtgttg accttatgt ggaagtta cgttagact ttgtctgaa  
gtgggtatc atgtgtgta agagcgcg catgatga cggatttcgc tccagcgt  
aagcagcat ggaacccgc gatttgata ctgtgttg gattcttg agtcgtgg  
ttgtccga tgaagcga aaagcggg aaagtcca atgtatga ttgtgtgt  
aagtagcga agaatgggt cagactga accatacag cagttttg cccgagtg  
caagacagt atcgtgga gtttatat ccttgaccg tctaaactg tggagtctc  
gaacttga ggaattctc tgaagctcc gatgatgac ctgacacag gatggagag  
atacagatc ggtgtgag acaaggtg acacaattc atatcctcg  
ttgtttgt taccatcgg tatgaanaa atgggtgaa ttgaagtcg agtcggtt  
gcatgcctt ctgtgtcc tgaattgt gaactatg gaacgcgt tctgctcgg  
atgcactaca taagcctct aggtgtaga caacagatg catlaaggg ggcgcgac  
aaatggtg cagcttggt gctcagga gaaccacat tggaccaga gtagttcga  
tatagttag atcacatgc atcgtatg aacatgga aaagcaagc gaattgtac  
agaattgtg gttgttag tggcagtg ggttagta agtggttga taatatcag  
cgttgagga atccagatc gacgtgta gtccatc ttatctggt tctgtttg  
taccgtgatt tggatccc gactgattc ttgtacgtg tgaatcgg agttgtgac  
taccgttta gaaccaat accgctggt atggatatc gttatcaca agtgaaac  
gtcgtatc atgagtga tgaagattc gacacatc caagctcaag ggcgcagaa  
gtaaatcag ctagttaga ccatlaag atcttagag tgaagttca gaccattca  
ggaatttg cagcgaag agacagat caagcttg tgaactgag agatccgga  
gcgaacag tttctagc atctgttg gtaacacaa tggttcgtg tgcagttct  
gcgaataat ggcgtgag tcatagatg tgaatagtg tgcacagg gaacaagc  
acgaagaat ctctaaga atcttca tgcctgct tgaatttc ttcaatgct  
gtcgtcgg ctcacatc gtctccggc gctcagtc agacaaatt tctctcaat  
cccatctc ccgcttgc ccgtcttc tccacggga aatcaccgc aactctcc  
gtctagga tggctccc gaanaagtg acaaatatg atgcagtg gaagaaca  
tggtagcag ctgattgt ttgcaagg agtagcaaa taacgttga tgtttcaag  
aagtcgga agcgaagt gttgacac gttgagaat ctggctgct gtcanaaga  
gaggggttg gactcacat gcatcttt gagaagctta aagtcttc caagcagag  
gactctgct tctcagct ccttgagac tatgtgaa catcgtcgc gttctagcc  
tggctcat taccagctct cagcgtct atttagcgc tgggtgat ccgagtagc  
tcaactact tagtggttg tggcggtt tggcgggt ccttgctg tccaggggtt  
gtttgttg ttgttctg ttgtttgat ggaactcag aagtcagtg attcttctc  
tgtaaccca acataaacc atgtctgc caattgatt ttgtcagtg ctgatttata  
gctgtatggt tcaagttgtt atggttagtc caagacataa gctgagtag agaaagaagc  
tttataataa ttaacaaat att

<210> 401

<211> 1276

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..1276:Ceres Seq. ID 2028730

<400>401

Met Asn Thr Thr Pro Phe His Ser Asp Pro Pro Pro Ser Arg Ile Gln  
1 5 10 15  
Arg Lys Leu Val Val Glu Val Val Glu Ala Arg Asn Ile Leu Pro Lys  
20 25 30  
Asp Gly Gln Gly Ser Ser Ala Tyr Val Val Val Asp Phe Asp Ala  
35 40 45

Gln Lys Lys Arg Thr Ser Thr Lys Phe Arg Asp Leu Asn Pro Ile Trp  
50 55 60  
Asn Glu Met Leu Asp Phe Ala Val Ser Asp Pro Lys Asn Met Asp Tyr  
65 70 75 80  
Asp Glu Leu Asp Ile Glu Val Tyr Asn Asp Lys Arg Phe Gly Asn Gly  
85 90 95  
Gly Gly Arg Lys Asn His Phe Leu Gly Arg Val Lys Ile Tyr Gly Ser  
100 105 110  
Gln Phe Ser Arg Arg Gly Glu Gly Leu Val Tyr Phe Pro Leu Glu  
115 120 125  
Lys Lys Ser Val Phe Ser Trp Ile Arg Gly Glu Ile Gly Leu Lys Ile  
130 135 140  
Tyr Tyr Tyr Asp Glu Ala Ala Asp Glu Asp Thr Ala Gly Gly Gly  
145 150 155 160  
Gly Gln Gln Gln Gln Gln Gln Gln Phe His Pro Pro Gln Gln  
165 170 175  
Glu Ala Asp Glu Gln Gln His Gln Gln Phe His Pro Pro Gln  
180 185 190  
Gln Met Met Asn Ile Pro Pro Glu Lys Pro Asn Val Val Val Glu  
195 200 205  
Glu Gly Arg Val Phe Glu Ser Ala Gln Ser Gln Arg Tyr Thr Glu Thr  
210 215 220  
His Gln Gln Pro Pro Val Val Ile Val Glu Ser Pro Pro Gln His  
225 230 235 240  
Val Met Gln Gly Pro Asn Asp Asn His Pro His Arg Asn Asp Asn His  
245 250 255  
Pro Gln Arg Pro Pro Ser Pro Pro Pro Ser Ala Gly Glu Val  
260 265 270  
His Tyr Tyr Pro Pro Glu Val Arg Lys Met Gln Val Gly Arg Pro  
275 280 285  
Gly Gly Asp Arg Ile Arg Val Thr Lys Arg Pro Pro Asn Gly Asp Tyr  
290 295 300  
Ser Pro Arg Val Ile Asn Ser Lys Thr Gly Gly Glu Thr Thr Met  
305 310 315 320  
Glu Lys Lys Thr His His Pro Tyr Asn Leu Val Glu Pro Met Gln Tyr  
330 335 340  
Leu Phe Val Arg Ile Val Lys Ala Arg Gly Leu Pro Pro Asn Glu Ser  
345 350  
Ala Tyr Val Lys Val Arg Thr Ser Asn His Phe Val Arg Ser Lys Pro  
355 360 365  
Ala Val Asn Arg Pro Gly Glu Ser Val Asp Ser Pro Glu Trp Asn Gln  
370 375 380  
Val Phe Ala Leu Gly His Asn Arg Ser Asp Ser Ala Val Thr Gly Ala  
385 390 395 400  
Thr Leu Glu Ile Ser Ala Trp Asp Ala Ser Ser Glu Ser Phe Leu Gly  
405 410 415  
Gly Val Cys Phe Asp Leu Ser Glu Val Pro Val Arg Asp Pro Pro Asp  
420 425 430  
Ser Pro Leu Ala Pro Gln Trp Tyr Arg Leu Glu Gly Ser Gly Ala Asp  
435 440 445  
Gln Asn Ser Gly Arg Ile Ser Gly Asp Ile Gln Leu Ser Val Trp Ile  
450 455 460  
Gly Thr Gln Val Asp Glu Ala Phe Pro Glu Ala Trp Ser Ser Asp Ala  
465 470 475 480  
Pro His Val Ala His Thr Arg Ser Lys Val Tyr Gln Ser Pro Lys Leu  
485 490 495  
Trp Tyr Leu Arg Val Thr Val Leu Glu Ala Gln Asp Leu His Ile Ala  
500 505 510  
Pro Asn Leu Pro Pro Leu Thr Ala Pro Glu Ile Arg Val Lys Ala Gln  
515 520 525  
Leu Gly Phe Gln Ser Ala Arg Thr Arg Arg Gly Ser Met Asn Asn His



530 Ser Gly Ser Phe His Trp His Glu Asp Met Ile Phe Val Ala Gly Glu 540  
 545 Pro Leu Glu Asp Cys Leu Val Leu Met Val Glu Asp Arg Thr Thr Lys 550  
 555 Glu Ala Thr Leu Leu Gly His Ala Met Ile Pro Val Ser Ser Ile Glu 560  
 565 Gln Arg Ile Asp Glu Arg Phe Val Pro Ser Lys Trp His Thr Leu Glu 570  
 580 Gly Glu Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly 585  
 590 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly 595  
 600 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly 605  
 610 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly 615  
 620 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly 625  
 630 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly 635  
 640 Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly 645  
 650 Arg Pro Thr Ala Lys Gln Leu Trp Lys Pro Pro Ile Gly Ile Leu Glu 655  
 660 Leu Gly Ile Leu Gly Ala Arg Gly Leu Leu Pro Met Lys Ala Lys Asn 665  
 670 Gly Gly Lys Gly Ser Thr Asp Ala Tyr Cys Val Ala Lys Tyr Gly Lys 675  
 680 Lys Trp Val Arg Thr Arg Thr Ile Thr Asp Ser Phe Asp Pro Arg Trp 685  
 690 His Glu Gln Tyr Thr Trp Gln Val Tyr Asp Pro Cys Thr Val Leu Thr 700  
 705 Val Gly Val Phe Asp Asn Trp Arg Met Phe Ser Asp Ala Ser Asp Asp 710  
 715 Arg Pro Asp Thr Arg Ile Gly Lys Ile Arg Ile Arg Val Ser Thr Leu 715  
 720 Glu Ser Asn Lys Val Tyr Thr Asn Ser Tyr Pro Leu Leu Val Leu Leu 725  
 730 Pro Ser Gly Met Lys Lys Met Gly Glu Ile Glu Val Ala Val Arg Phe 735  
 740 Ala Cys Pro Ser Leu Leu Pro Asp Val Cys Ala Ala Tyr Gly Gln Pro 745  
 750 Leu Leu Pro Arg Met His Tyr Ile Arg Pro Leu Gly Val Ala Gln Gln 755  
 760 Asp Ala Leu Arg Gly Ala Ala Thr Lys Lys Met Val Ala Ala Trp Leu Ala 765  
 770 Arg Ala Glu Pro Pro Leu Gly Pro Glu Val Val Arg Tyr Met Leu Asp 775  
 780 Ala Asp Ser His Ala Trp Ser Met Arg Lys Ser Lys Ala Asn Trp Tyr 785  
 790 Arg Ile Val Gly Val Leu Ala Trp Ala Val Gly Leu Ala Lys Trp Leu 795  
 800 Asp Asn Ile Arg Arg Trp Arg Asn Pro Val Thr Thr Val Leu Val His 805  
 810 Ile Leu Tyr Leu Val Leu Val Trp Tyr Trp Pro Asp Leu Val Pro Thr 815  
 820 Ala Phe Leu Tyr Val Val Met Ile Gly Val Trp Tyr Arg Phe Arg 825  
 830 Pro Lys Ile Pro Ala Gly Met Asp Ile Arg Leu Ser Gln Ala Glu Thr 835  
 840 Val Asp Pro Asp Glu Leu Asp Glu Glu Phe Asp Thr Ile Pro Ser Ser 845  
 850 Arg Arg Pro Glu Val Ile Arg Ala Arg Tyr Asp Arg Leu Arg Ile Leu 855  
 860 Ala Val Arg Val Gln Thr Ile Leu Gly Asp Phe Ala Ala Gln Gly Glu 865  
 870 Arg Ile Gln Ala Leu Val Ser Trp Arg Asp Pro Arg Ala Thr Lys Leu 875  
 880 1010

Phe Ile Ala Ile Cys Leu Val Ile Thr Ile Val Leu Tyr Ala Val Pro 1025  
 1030 Ala Lys Met Val Ala Val Ala Leu Gly Val Ser Asp Ser Val Pro Thr 1035  
 1040 Ala Lys Gln Asp Thr Lys Glu Ser Leu Lys Lys Ser Phe Ser Ser Leu 1045  
 1050 Arg Phe Asp Phe Ser Ser Met Ala Val Val Gly Ala Pro Ile Ser Ser 1055  
 1060 Pro Ala Ala Gln Leu Gln Thr Gln Phe Leu Ser Asn Pro Ile Leu Pro 1065  
 1070 Arg Phe Arg Arg Ser Phe Thr Gly Lys Ser Pro Ala Thr Phe Ser 1075  
 1080 Val Val Ala Met Ala Pro Gln Lys Lys Val Asn Lys Tyr Asp Ala Lys 1085  
 1090 Trp Lys Lys Gln Trp Tyr Gly Ala Gly Leu Phe Phe Glu Gly Ser Glu 1095  
 1100 Gln Ile Asn Val Asp Val Phe Lys Lys Leu Glu Lys Arg Lys Val Leu 1105  
 1110 Ser Asn Val Glu Lys Ser Gly Leu Leu Ser Lys Ala Glu Gly Leu Gly 1115  
 1120 Leu Thr Leu Ser Ser Leu Glu Lys Leu Lys Val Phe Ser Lys Ala Glu 1125  
 1130 Asp Leu Gly Leu Leu Ser Leu Leu Glu Asn Leu Ala Gly Thr Ser Pro 1135  
 1140 Ala Val Leu Ala Ser Ala Ala Leu Pro Ala Leu Thr Ala Ile Val 1145  
 1150 Ala Val Val Leu Ile Pro Asp Asp Ser Thr Thr Leu Val Val Ala Gln 1155  
 1160 Ala Val Leu Ala Gly Ala Leu Ala Leu Thr Gly Val Val Leu Leu Val 1165  
 1170 Gly Ser Val Val Leu Asp Gly Leu Gln Glu Ala Asp 1175  
 1180 1265  
 1270  
 <210> 402  
 <211> 1210  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..1210:Ceres Seq. ID 2028731  
 <400>402  
 Met Leu Asp Phe Ala Val Ser Asp Pro Lys Asn Met Asp Tyr Asp Glu 1  
 5 Leu Asp Ile Glu Val Tyr Asn Asp Lys Arg Phe Gly Asn Gly Gly Gly 15  
 20 Arg Lys Asn His Phe Leu Gly Arg Val Lys Ile Tyr Gly Ser Gln Phe 30  
 35 Ser Arg Arg Gly Glu Glu Gly Leu Val Tyr Phe Pro Leu Glu Lys Lys 45  
 50 Ser Val Phe Ser Trp Ile Arg Gly Glu Ile Gly Leu Lys Ile Tyr Tyr 60  
 65 Tyr Asp Glu Ala Ala Asp Glu Asp Thr Ala Gly Gly Gly Gly Gln 75  
 80 Gln Gln Gln Gln Gln Gln Gln Phe His Pro Pro Gln Gln Glu Ala 85  
 90 Asp Glu Gln Gln His Gln Gln Gln Phe His Pro Pro Gln Gln Met 95  
 100 Met Asn Ile Pro Pro Glu Lys Pro Asn Val Val Val Glu Glu Gly 105  
 110 125

130 Arg Val Phe Glu Ser Ala Gln Ser Gln Arg Tyr Thr Glu Thr His Gln 140  
 145 Gln Pro Pro Val Val Ile Val Glu Glu Ser Pro Pro Gln His Val Met 155  
 165 Gln Gly Pro Asn Asp Asn His Pro His Arg Asn Asp Asn His Pro Gln 175  
 180 Arg Pro Pro Ser Pro Pro Pro Pro Ala Gly Glu Val His Tyr 190  
 195 Tyr Pro Pro Glu Val Arg Lys Met Gln Val Gly Arg Pro Pro Gly Gly 205  
 210 Asp Arg Ile Arg Val Thr Lys Arg Pro Pro Asn Gly Asp Tyr Ser Pro 220  
 225 Arg Val Ile Asn Ser Lys Thr Gly Gly Gly Thr Thr Met Glu Lys 235  
 245 Lys Thr His His Pro Tyr Asn Leu Val Glu Pro Met Gln Tyr Leu Phe 255  
 260 Val Arg Ile Val Lys Ala Arg Gly Leu Pro Pro Asn Glu Ser Ala Tyr 270  
 275 Val Lys Val Arg Thr Ser Asn His Phe Val Arg Ser Lys Pro Ala Val 285  
 290 Asn Arg Pro Gly Glu Ser Val Asp Ser Pro Glu Tyr Asn Gln Val Phe 300  
 305 Ala Leu Gly His Asn Arg Ser Asp Ser Ala Val Thr Gly Ala Thr Leu 315  
 325 Glu Ile Ser Ala Trp Asp Ala Ser Ser Glu Ser Phe Leu Gly Gly Val 335  
 340 Cys Phe Asp Leu Ser Glu Val Pro Val Arg Asp Pro Pro Asp Ser Pro 350  
 355 Leu Ala Pro Gln Trp Tyr Arg Leu Glu Gly Ser Gly Ala Asp Gln Asn 365  
 370 Ser Gly Arg Ile Ser Gly Asp Ile Gln Leu Ser Val Trp Ile Gly Thr 380  
 385 Gln Val Asp Glu Ala Phe Pro Glu Ala Trp Ser Ser Asp Ala Pro His 395  
 405 Val Ala His Thr Arg Ser Lys Val Tyr Gln Ser Pro Lys Leu Trp Tyr 415  
 420 Leu Arg Val Thr Val Leu Glu Ala Gln Asp Leu His Ile Ala Pro Asn 430  
 435 Leu Pro Pro Leu Thr Ala Pro Glu Ile Arg Val Lys Ala Gln Leu Gly 445  
 450 Phe Gln Ser Ala Arg Thr Arg Arg Gly Ser Met Asn Asn His Ser Gly 460  
 465 Ser Phe His Trp His Glu Asp Met Ile Phe Val Ala Gly Glu Pro Leu 475  
 485 Glu Asp Cys Leu Val Leu Met Val Glu Asp Arg Thr Thr Lys Glu Ala 495  
 500 Thr Leu Leu Gly His Ala Met Ile Pro Val Ser Ser Ile Glu Gln Arg 510  
 515 Ile Asp Glu Arg Phe Val Pro Ser Lys Trp His Thr Leu Glu Gly Glu 525  
 530 Gly Gly Gly Gly Gly Gly Gly Gly Pro Gly Gly Gly Gly Gly Gly 540  
 545 Gly Pro Tyr Cys Gly Arg Ile Ser Leu Arg Leu Cys Leu Glu Gly Gly 555  
 565 Tyr His Val Leu Glu Glu Ala Ala His Val Cys Ser Asp Phe Arg Pro 575  
 580 Thr Ala Lys Gln Leu Trp Lys Pro Pro Ile Gly Ile Leu Glu Leu Gly 590  
 595 Ile Leu Gly Ala Arg Gly Leu Leu Pro Met Lys Ala Lys Asn Gly Gly 605  
 610

Lys Gly Ser Thr Asp Ala Tyr Cys Val Ala Lys Tyr Gly Lys Lys Trp 625  
 630 Val Arg Thr Arg Thr Ile Thr Asp Ser Phe Asp pro Arg Trp His Glu 640  
 645 Gln Tyr Thr Trp Gln Val Tyr Asp Pro Cys Thr Val Leu Thr Val Gly 655  
 660 Val Phe Asp Asn Trp Arg Met Phe Ser Asp Ala Ser Asp Asp Arg Pro 670  
 675 Asp Thr Arg Ile Gly Lys Ile Arg Ile Arg Val Ser Thr Leu Glu Ser 685  
 690 Asn Lys Val Tyr Thr Asn Ser Tyr Pro Leu Leu Val Leu Leu Pro Ser 700  
 705 Gly Met Lys Lys Met Gly Glu Ile Glu Val Ala Val Arg Phe Ala Cys 715  
 720 Pro Ser Leu Leu Pro Asp Val Cys Ala Ala Tyr Gly Gln Pro Leu Leu 730  
 735 Pro Arg Met His Tyr Ile Arg Pro Leu Gly Val Ala Gln Gln Asp Ala 745  
 750 Leu Arg Gly Ala Ala Thr Lys Met Val Ala Ala Trp Leu Ala Arg Ala 765  
 770 Glu Pro Pro Leu Gly Pro Glu Val Val Arg Tyr Met Leu Asp Ala Asp 780  
 785 Ser His Ala Trp Ser Met Arg Lys Ser Lys Ala Asn Trp Tyr Arg Ile 815  
 820 Val Gly Val Leu Ala Trp Ala Val Gly Leu Ala Lys Trp Leu Asp Asn 830  
 835 Ile Arg Arg Trp Arg Asn Pro Val Thr Thr Val Leu Val His Ile Leu 845  
 850 Tyr Leu Val Leu Val Trp Tyr Pro Asp Leu Val Val Pro Thr Ala Phe 860  
 865 Leu Tyr Val Val Met Ile Gly Val Trp Tyr Arg Phe Arg Pro Lys 875  
 880 Ile Pro Ala Gly Met Asp Ile Arg Leu Ser Gln Ala Glu Thr Val Asp 895  
 895 Pro Asp Glu Leu Asp Glu Glu Phe Asp Thr Ile Pro Ser Ser Arg Arg 910  
 915 Pro Glu Val Ile Arg Ala Arg Tyr Asp Arg Leu Arg Ile Leu Ala Val 925  
 930 Arg Val Gln Thr Ile Leu Gly Asp Phe Ala Ala Gln Gly Glu Arg Ile 940  
 945 Gln Ala Leu Val Ser Trp Arg Asp Pro Arg Ala Thr Lys Leu Phe Ile 955  
 960 Ala Ile Cys Leu Val Ile Thr Ile Val Leu Tyr Ala Val Pro Ala Lys 975  
 980 Met Val Ala Val Ala Leu Gly Val Ser Asp Ser Val Pro Thr Ala Lys 990  
 995 Gln Asp Thr Lys Glu Ser Leu Lys Lys Ser Phe Ser Ser Leu Arg Phe 1005  
 1010 Asp Phe Ser Ser Met Ala Val Val Gly Ala Pro Ile Ser Ser Pro Ala 1020  
 1025 Ala Gln Leu Gln Thr Gln Phe Leu Ser Asn Pro Ile Leu Pro Arg Phe 1035  
 1040 Arg Arg Ser Phe Ser Thr Gly Lys Ser Pro Ala Thr Phe Ser Val Val 1055  
 1060 Ala Met Ala Pro Gln Lys Lys Val Asn Lys Tyr Asp Ala Lys Trp Lys 1070  
 1080 Lys Gln Trp Tyr Gly Ala Gly Leu Phe Phe Glu Gly Ser Glu Gln Ile 1085  
 1090 Asn Val Asp Val Phe Lys Lys Leu Glu Lys Arg Lys Val Leu Ser Asn 1100  
 1105 Val Glu Lys Ser Gly Leu Leu Ser Lys Ala Glu Gly Leu Gly Leu Thr

1105 1110 1115 1120  
 Leu Ser Ser Leu Glu Lys Leu Lys Val Phe Ser Lys Ala Glu Asp Leu  
 1125 1130 1135  
 Gly Leu Ser Leu Leu Glu Asn Leu Ala Gly Thr Ser Pro Ala Val  
 1140 1145 1150  
 Leu Ala Ser Ala Ala Leu Pro Ala Leu Thr Ala Ala Ile Val Ala Val  
 1155 1160 1165  
 Val Leu Ile Pro Asp Asp Ser Thr Thr Leu Val Val Ala Gln Ala Val  
 1170 1175 1180  
 Leu Ala Gly Ala Leu Ala Leu Thr Gly Val Val Leu Leu Val Gly Ser  
 1185 1190 1195 1200  
 Val Val Leu Asp Gly Leu Gln Glu Ala Asp  
 1205 1210

&lt;210&gt; 403

&lt;211&gt; 1199

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..1199:Ceres Seq. ID 2028732

&lt;400&gt; 403

Met Asp Tyr Asp Glu Leu Asp Ile Glu Val Tyr Asn Asp Lys Arg Phe  
 1 5 10 15  
 Gly Asn Gly Gly Arg Lys Asn His Phe Leu Gly Arg Val Lys Ile  
 20 25 30  
 Tyr Gly Ser Gln Phe Ser Arg Arg Gly Glu Gly Leu Val Tyr Phe  
 35 40 45  
 Pro Leu Glu Lys Lys Ser Val Phe Ser Trp Ile Arg Gly Glu Ile Gly  
 50 55 60  
 Leu Lys Ile Tyr Tyr Asp Glu Ala Ala Asp Glu Asp Thr Ala Gly  
 65 70 75  
 Gly Gly Gly Gln Gln Gln Gln Gln Gln Gln Phe His Pro  
 80 85 90 95  
 Pro Gln Gln Glu Ala Asp Glu Gln Gln His Gln Gln Phe His Pro  
 100 105 110  
 Pro Pro Gln Gln Met Met Asn Ile Pro Pro Glu Lys Pro Asn Val Val  
 115 120 125  
 Val Val Glu Glu Gly Arg Val Phe Glu Ser Ala Gln Ser Gln Arg Tyr  
 130 135 140  
 Thr Glu Thr His Gln Gln Pro Pro Val Ile Val Glu Glu Ser Pro  
 145 150 155 160  
 Pro Gln His Val Met Gln Gly Pro Asn Asp Asn His Pro His Arg Asn  
 165 170 175  
 Asp Asn His Pro Gln Arg Pro Pro Ser Pro Pro Pro Ser Ala  
 180 185 190  
 Gly Glu Val His Tyr Tyr Pro Pro Glu Val Arg Lys Met Gln Val Gly  
 195 200 205  
 Arg Pro Pro Gly Gly Asp Arg Ile Arg Val Thr Lys Arg Pro Pro Asn  
 210 215 220  
 Gly Asp Tyr Ser Pro Arg Val Ile Asn Ser Lys Thr Gly Gly Glu  
 225 230 235 240  
 Thr Thr Met Glu Lys Lys Thr His His Pro Tyr Asn Leu Val Glu Pro  
 245 250 255  
 Met Gln Tyr Leu Phe Val Arg Ile Val Lys Ala Arg Gly Leu Pro Pro  
 260 265 270  
 Asn Glu Ser Ala Tyr Val Lys Val Arg Thr Ser Asn His Phe Val Arg  
 275 280 285

Ser Lys Pro Ala Val Asn Arg Pro Gly Glu Ser Val Asp Ser Pro Glu  
 290 295 300  
 Trp Asn Gln Val Phe Ala Leu Gly His Asn Arg Ser Asp Ser Ala Val  
 305 310 315 320  
 Thr Gly Ala Thr Leu Glu Ile Ser Ala Trp Asp Ala Ser Ser Glu Ser  
 325 330 335 340  
 Phe Leu Gly Gly Val Cys Phe Asp Leu Ser Glu Val Pro Val Arg Asp  
 345 350  
 Pro Pro Asp Ser Pro Leu Ala Pro Gln Trp Tyr Arg Leu Gly Ser  
 355 360 365  
 Gly Ala Asp Gln Asn Ser Gly Arg Ile Ser Gly Asp Ile Gln Leu Ser  
 370 375 380  
 Val Trp Ile Gly Thr Gln Val Asp Glu Ala Phe Pro Glu Ala Trp Ser  
 385 390 395 400  
 Ser Asp Ala Pro His Val Ala His Thr Arg Ser Lys Val Tyr Gln Ser  
 405 410 415  
 Pro Lys Leu Trp Tyr Leu Arg Val Thr Val Leu Glu Ala Gln Asp Leu  
 420 425 430  
 His Ile Ala Pro Asn Leu Pro Pro Leu Thr Ala Pro Glu Ile Arg Val  
 435 440 445  
 Lys Ala Gln Leu Gly Phe Gln Ser Ala Arg Thr Arg Arg Gly Ser Met  
 450 455 460  
 Asn Asn His Ser Gly Ser Phe His Trp His Glu Asp Met Ile Phe Val  
 465 470 475 480  
 Ala Gly Glu Pro Leu Glu Asp Cys Leu Val Leu Met Val Glu Asp Arg  
 485 490 495  
 Thr Thr Lys Glu Ala Thr Leu Leu Gly His Ala Met Ile Pro Val Ser  
 500 505 510  
 Ser Ile Glu Gln Arg Ile Asp Glu Arg Phe Val Pro Ser Lys Trp His  
 515 520 525  
 Thr Leu Glu Gly Glu Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly  
 530 535 540  
 Gly Gly Gly Gly Gly Gly Pro Tyr Cys Gly Arg Ile Ser Leu Arg Leu  
 545 550 555 560  
 Cys Leu Glu Gly Gly Tyr His Val Leu Glu Glu Ala Ala His Val Cys  
 565 570 575  
 Ser Asp Phe Arg Pro Thr Ala Lys Gln Leu Trp Lys Pro Ile Gly  
 580 585 590  
 Ile Leu Glu Leu Gly Ile Leu Gly Ala Arg Gly Leu Leu Pro Met Lys  
 595 600 605  
 Ala Lys Asn Gly Gly Lys Gly Ser Thr Asp Ala Tyr Cys Val Ala Lys  
 610 615 620  
 Tyr Gly Lys Lys Trp Val Arg Thr Arg Thr Ile Thr Asp Ser Phe Asp  
 625 630 635 640  
 Pro Arg Trp His Glu Gln Tyr Thr Trp Gln Val Tyr Asp Pro Cys Thr  
 645 650 655  
 Val Leu Thr Val Gly Val Phe Asp Asn Trp Arg Met Phe Ser Asp Ala  
 660 665 670  
 Ser Asp Asp Arg Pro Asp Thr Arg Ile Gly Lys Ile Arg Ile Arg Val  
 675 680 685  
 Ser Thr Leu Glu Ser Asn Lys Val Tyr Thr Asn Ser Tyr Pro Leu Leu  
 690 695 700  
 Val Leu Leu Pro Ser Gly Met Lys Lys Met Gly Glu Ile Glu Val Ala  
 705 710 715 720  
 Val Arg Phe Ala Cys Pro Ser Leu Leu Pro Asp Val Cys Ala Ala Tyr  
 725 730 735  
 Gly Gln Pro Leu Leu Pro Arg Met His Tyr Ile Arg Pro Leu Gly Val  
 740 745 750  
 Ala Gln Gln Asp Ala Leu Arg Gly Ala Ala Thr Lys Met Val Ala Ala  
 755 760 765  
 Trp Leu Ala Arg Ala Glu Pro Pro Leu Gly Pro Glu Val Val Arg Tyr

770 Met Leu Asp Ala Asp Ser His Ala Trp Ser Met Arg Lys Ser Lys Ala 780  
 785 Asn Trp Tyr Arg Ile Val Gly Val Leu Ala Trp Ala Val Gly Leu Ala 800  
 805 Lys Trp Leu Asp Asn Ile Arg Arg Trp Arg Asn Pro Val Thr Thr Val 815  
 820 Leu Val His Ile Leu Tyr Leu Val Leu Val Trp Tyr Pro Asp Leu Val 830  
 835 Val Pro Thr Ala Phe Leu Tyr Val Val Met Ile Gly Val Trp Tyr Tyr 845  
 850 Arg Phe Arg Pro Lys Ile Pro Ala Gly Met Asp Ile Arg Leu Ser Gln 860  
 865 Ala Glu Thr Val Asp Pro Asp Glu Leu Asp Glu Glu Phe Asp Thr Ile 880  
 885 Pro Ser Ser Arg Arg Pro Glu Val Ile Arg Ala Arg Tyr Asp Arg Leu 895  
 900 Arg Ile Leu Ala Val Arg Val Gln Thr Ile Leu Gly Asp Phe Ala Ala 910  
 915 Gln Gly Glu Arg Ile Gln Ala Leu Val Ser Trp Arg Asp Pro Arg Ala 925  
 930 Thr Lys Leu Phe Ile Ala Ile Cys Leu Val Ile Thr Ile Val Leu Tyr 940  
 945 Ala Val Pro Ala Lys Met Val Ala Val Ala Leu Gly Val Ser Asp Ser 955  
 965 Val Pro Thr Ala Lys Gln Asp Thr Lys Glu Ser Leu Lys Lys Ser Phe 975  
 980 Ser Ser Leu Arg Phe Asp Phe Ser Ser Met Ala Val Val Gly Ala Pro 985  
 995 Ile Ser Ser Pro Ala Ala Gln Leu Gln Thr Gln Phe Leu Ser Asn Pro 1000  
 1010 Ile Leu Pro Arg Phe Arg Arg Ser Phe Ser Thr Gly Lys Ser Pro Ala 1020  
 1025 Thr Phe Ser Val Val Ala Met Ala Pro Gln Lys Lys Val Asn Lys Tyr 1035  
 1040 Asp Ala Lys Trp Lys Lys Gln Trp Tyr Gly Ala Gly Leu Phe Phe Glu 1055  
 1060 Gly Ser Glu Gln Ile Asn Val Asp Val Phe Lys Lys Leu Glu Lys Arg 1070  
 1075 Lys Val Leu Ser Asn Val Glu Lys Ser Gly Leu Leu Ser Lys Ala Glu 1085  
 1090 Gly Leu Gly Leu Thr Leu Ser Ser Leu Glu Lys Leu Lys Val Phe Ser 1100  
 1105 Lys Ala Glu Asp Leu Gly Leu Ser Leu Leu Glu Asn Leu Ala Gly 1115  
 1120 Thr Ser Pro Ala Val Leu Ala Ser Ala Ala Leu Pro Ala Leu Thr Ala 1130  
 1135 Ala Ile Val Ala Val Val Leu Ile Pro Asp Asp Ser Thr Thr Leu Val 1145  
 1150 Val Ala Gln Ala Val Leu Ala Gly Ala Leu Ala Leu Thr Gly Val Val 1165  
 1170 Leu Leu Val Gly Ser Val Val Leu Asp Gly Leu Glu Ala Ala Asp 1180  
 1185

&lt;210&gt; 404

&lt;211&gt; 528

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..528:Ceres Seq. ID 2029079

&lt;400&gt;404

aaaaaacta cttcaacttg cacaagaata agaaacccat gctgtgtca gcaatcttg 60  
 gtaccggaat cgtccacgtc gctgctctc cgtttctccg ccaatttcaa gtccaaaat 120  
 tgggtaatg aggtgattg gggatgtga ttgagtgtc gtccaggcca cagaagaat 180  
 cgcacactga tcaaggagc agcaggcca agaagactca gcttgggac attaagaga 240  
 agctactgt gtatgtcct cttctcctc ttccgggga atggagtccg ttactcttg 300  
 ctccaaga cagtggtgc gcggttgct cttctcctc cggagattg gttcaggt 360  
 ctgcctagt atagagcta tctgtggtt tggttgaat ctgattgag agctttgtt 420  
 ggtgttgtt ttgttaagt ttctcagtt ttctcgtc tgtttctat tatccggtt 480  
 atgttggtt ggtgaatgaa tgtcaatgga ccaaacgtca atttttt 528

&lt;210&gt; 405

&lt;211&gt; 121

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..121:Ceres Seq. ID 2029080

&lt;400&gt;405

Lys Thr His Phe His Leu His Lys Asp Lys Glu Thr Met Ser Val Ser 1  
 5 Ala Ile Phe Gly Thr Gly Ile Val Thr Val Ala Ala Ser Pro Val Leu 10  
 20 Arg Gln Phe Gln Val Pro Lys Leu Gly Asn Gly Gly Leu Gly Met 25  
 35 Val Ile Glu Cys Ser Ser Arg Pro Gln Lys Lys Ser Thr Ala His His 40  
 50 Arg Lys Thr Arg Pro Lys Lys Thr Gln Pro Trp Asp Ile Lys Arg Lys 55  
 65 Pro Thr Val Tyr Ala Pro Leu Pro Pro Leu Pro Ala Glu Trp Ser Pro 70  
 85 Phe Thr Leu Ala Ser Asn Asp Gly Gly Ala Val Ala Val Ala Ser Pro 85  
 100 Ala Gly Asp Leu Val Ser Gly Ser Ala 105  
 115

&lt;210&gt; 406

&lt;211&gt; 109

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..109:Ceres Seq. ID 2029081

&lt;400&gt;406

Met Ser Val Ser Ala Ile Phe Phe Gly Thr Gly Ile Val Thr Val Ala Ala 1  
 5 Ser Pro Val Leu Arg Gln Phe Gln Val Pro Lys Leu Gly Asn Gly Gly 10  
 20 Gly Leu Gly Met Val Ile Glu Cys Ser Ser Arg Pro Gln Lys Lys Ser 25  
 35 Thr Ala His His Arg Lys Thr Arg Pro Lys Lys Thr Gln Pro Trp Asp 40  
 55

Ile Lys Arg Lys Pro Thr Val Tyr Ala Pro Leu Pro Leu Pro Ala  
65 70 75 80  
Glu Trp Ser Pro Phe Thr Leu Ala Ser Asn Asp Gly Gly Ala Ala Val  
85 90 95  
Ala Ala Ser Pro Ala Gly Asp Leu Val Ser Gly Ser Ala  
100 105

&lt;210&gt; 407

&lt;211&gt; 74

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..74:Ceres Seq. ID 2029082

&lt;400&gt;407

Met Val Ile Glu Cys Ser Ser Arg Pro Gln Lys Lys Ser Thr Ala His  
1 5 10 15  
His Arg Lys Thr Arg Pro Lys Lys Thr Gln Pro Trp Asp Ile Lys Arg  
20 25 30  
Lys Pro Thr Val Tyr Ala Pro Leu Pro Pro Leu Pro Ala Glu Trp Ser  
35 40 45  
Pro Phe Thr Leu Ala Ser Asn Asp Gly Gly Ala Val Ala Ala Ser  
50 55 60  
Pro Ala Gly Asp Leu Val Ser Gly Ser Ala  
65 70

&lt;210&gt; 408

&lt;211&gt; 787

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..787:Ceres Seq. ID 2029225

<400>408  
aggattctta ggccttctct tcattgtgca gacacagca cgttttaaat cttctcttc  
120  
ttcctcttcg atattcgca gtgtgtgat ttggggaaa actttgtgg caaagagcga  
180  
gaaatggagc ggagccgcta agaaatcgc ggaatgtgct ttaaaagctt caaggactat  
240  
cgattgggat ggtatggcta agtctctgt cacagatgag gctcgtagag agttctctaa  
300  
ccttcgtgt gcttcgtag aggttaacac acagctccag accnaattta gtcaggaaac  
360  
tgaacctata gattgggat actataggaa ggtatttga gctgcatgt ttgacaagta  
420  
caaggagct tatgacagca ttgagattcc aaagtacgtt gacaagatta cttctgaata  
480  
caagccaag ttgatgctt tttgttggga actgaagaa gcagaacaga aatgcctcaa  
540  
gggtgtgaa cggttggaga agaaatgct tgaagtccaa gagatcagca aaaagctcag  
600  
cacactgat gcagatgagt actttgaga gaaccggaa ctcaaaaga agtttgaga  
660  
cgaaatccgt aatgaacact ggggatactg atcattgttc tccatctcg gcttggaaag  
720  
aaaactctct tttcttctct cgtctctct actgtgatt tgtgagccaa tcaataaat  
780  
ataagtaaca caattcaact aagcagtggt gagatcttca ttcaaggaa gataaacga  
787  
tttgggt

&lt;210&gt; 409

&lt;211&gt; 168

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..168:Ceres Seq. ID 2029226

&lt;400&gt;409

Met Scr Gly Ala Gly Lys Lys Ile Ala Asp Val Ala Phe Lys Ala Ser  
1 5 10 15  
Arg Thr Ile Asp Trp Asp Gly Met Ala Lys Val Leu Val Thr Asp Glu  
20 25 30  
Ala Arg Arg Glu Phe Ser Asn Leu Arg Arg Ala Phe Asp Glu Val Asn  
35 40 45  
Thr Gln Leu Gln Thr Lys Phe Ser Gln Glu Pro Glu Pro Ile Asp Trp  
50 55 60  
Asp Tyr Tyr Arg Lys Gly Ile Gly Ala Gly Ile Val Asp Lys Tyr Lys  
65 70 75 80  
Glu Ala Tyr Asp Ser Ile Glu Ile Pro Lys Tyr Val Asp Lys Val Thr  
85 90 95  
Pro Glu Tyr Lys Pro Lys Phe Asp Ala Leu Leu Val Glu Leu Lys Glu  
100 105 110  
Ala Glu Gln Lys Ser Leu Lys Glu Ser Glu Arg Leu Glu Lys Glu Ile  
115 120 125  
Ala Asp Val Gln Glu Ile Ser Lys Lys Leu Ser Thr Met Thr Ala Asp  
130 135 140  
Glu Tyr Phe Glu Lys His Pro Glu Leu Lys Lys Lys Phe Asp Asp Glu  
145 150 155 160  
Ile Arg Asn Asp Asn Trp Gly Tyr  
165

&lt;210&gt; 410

&lt;211&gt; 145

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..145:Ceres Seq. ID 2029227

&lt;400&gt;410

Met Ala Lys Val Leu Val Thr Asp Glu Ala Arg Arg Glu Phe Ser Asn  
1 5 10 15  
Leu Arg Arg Ala Phe Asp Glu Val Asn Thr Gln Leu Gln Thr Lys Phe  
20 25 30  
Ser Gln Glu Pro Glu Pro Ile Asp Trp Asp Tyr Tyr Arg Lys Gly Ile  
35 40 45  
Gly Ala Gly Ile Val Asp Lys Tyr Lys Glu Ala Tyr Asp Ser Ile Glu  
50 55 60  
Ile Pro Lys Tyr Val Asp Lys Val Thr Pro Glu Tyr Lys Pro Lys Phe  
65 70 75 80  
Asp Ala Leu Leu Val Glu Leu Lys Glu Ala Glu Gln Lys Ser Leu Lys  
85 90 95  
Glu Ser Glu Arg Leu Glu Lys Glu Ile Ala Asp Val Gln Glu Ile Ser  
100 105 110  
Lys Lys Leu Ser Thr Met Thr Ala Asp Glu Tyr Phe Glu Lys His Pro  
115 120 125  
Glu Leu Lys Lys Lys Phe Asp Asp Glu Ile Arg Asn Asp Asn Trp Gly  
130 135 140  
Tyr  
145

&lt;210&gt; 411



65 Phe Thr Arg Lys Ala Arg Thr Pro Arg Gly Lys Lys Ser His Thr 80  
 70 75  
 85 Met Val Phe His Asp Leu Val Pro Glu Met Ser Thr Glu Asp Gln Ala 95  
 100 105 110  
 115 Glu Ser Tyr Glu Val Glu Glu Gln Leu Ile Phe Glu Val Pro Val Met  
 120 125  
 130 Asn Ser Met Val Glu Glu Gln Cys Phe Asn Gln Ser Leu Glu Lys Gln  
 135 140  
 145 Asn Glu Phe Pro Met Met Pro Leu Ser Phe Lys Ser Ser Asp Glu Glu  
 150 155 160  
 165 Asp Asp Asn Ala Glu Ser Cys Leu Asn Gly Leu Phe Pro Thr Asp  
 170 175  
 180 Met Glu Leu Ala Gln Phe Thr Ala Asp Val Glu Thr Leu Leu Gly Gly  
 185 190  
 195 Gly Asp Arg Glu Phe His Ser Ile Glu Glu Leu Gly Leu Gly Glu Met  
 200 205  
 210 Leu Lys Ile Glu Lys Glu Val Glu Glu Glu Gly Val Val Thr Arg  
 215 220  
 225 Glu Val His Asp Gln Asp Glu Gly Asp Glu Thr Ser Pro Phe Glu Ile  
 230 235 240  
 245 Ser Phe Asp Tyr Thr Thr His Lys Thr Thr Phe Asp Glu Gly Glu  
 250 255  
 260 Glu Asp Glu Lys Glu Asp Val Met Lys Asn Val Met Glu Met Gly Val  
 265 270  
 275 Asn Glu Met Ser Gly Ile Lys Glu Glu Lys Lys Lys Ala Leu  
 280 285  
 290 Met Leu Arg Leu Asp Tyr Glu Ser Val Ile Ser Thr Thr Gly Gly Gln  
 295 300  
 305 Gly Ile Pro Thr Thr Ala Arg Val Pro Ser Glu Ile Asp Leu Asp Met  
 310 315 320  
 325 Val Cys Phe Pro Thr Thr Thr Met Gly Glu Ser Gly Ala Glu Ala His  
 330 335  
 340 His His Asn His Phe Arg Gly Leu Gly Leu His Leu Asp Ala Gly  
 345 350  
 355 Asp Gly Gly Arg Glu Ala Arg Val Ser Arg Tyr Arg Glu Lys Arg Arg  
 360 365  
 370 Thr Arg Leu Phe Ser Lys Lys Ile Arg Tyr Glu Val Arg Lys Leu Asn  
 375 380  
 385 Ala Glu Lys Arg Pro Arg Met Lys Lys Gly Arg Phe Val Lys Arg Ser  
 390 395 400  
 405 Ile Gly Val Ala His

&lt;210&gt; 414

&lt;211&gt; 309

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..309:Ceres Seq. ID 2029425

&lt;400&gt;414

Met Val Phe His Asp Leu Val Pro Glu Met Ser Thr Glu Asp Gln Ala  
 5 10 15  
 1 Glu Ser Tyr Glu Val Glu Glu Gln Leu Ile Phe Glu Val Pro Val Met  
 20 25 30  
 35 Asn Ser Met Val Glu Glu Gln Cys Phe Asn Gln Ser Leu Glu Lys Gln  
 40 45

Asn Glu Phe Pro Met Met Pro Leu Ser Phe Lys Ser Ser Asp Glu Glu  
 50 55 60  
 65 Asp Asp Asp Asn Ala Glu Ser Cys Leu Asn Gly Leu Phe Pro Thr Asp  
 70 75 80  
 85 Met Glu Leu Ala Gln Phe Thr Ala Asp Val Glu Thr Leu Leu Gly Gly  
 90 95  
 100 Gly Asp Arg Glu Phe His Ser Ile Glu Glu Leu Gly Leu Gly Glu Met  
 105 110  
 115 Leu Lys Ile Glu Lys Glu Val Glu Glu Glu Gly Val Val Thr Arg  
 120 125  
 130 Glu Val His Asp Gln Asp Glu Gly Asp Glu Thr Ser Pro Phe Glu Ile  
 135 140  
 145 Ser Phe Asp Tyr Glu Tyr Thr His Lys Thr Thr Phe Asp Glu Gly Glu  
 150 155 160  
 165 Glu Asp Glu Lys Glu Asp Val Met Lys Asn Val Met Glu Met Gly Val  
 170 175  
 180 Asn Glu Met Ser Gly Gly Ile Lys Glu Glu Lys Lys Glu Lys Ala Leu  
 185 190  
 195 Met Leu Arg Leu Asp Tyr Glu Ser Val Ile Ser Thr Thr Gly Gly Gln  
 200 205  
 210 Gly Ile Pro Thr Thr Ala Arg Val Pro Ser Glu Ile Asp Leu Asp Met  
 215 220  
 225 Val Cys Phe Pro Thr Thr Thr Met Gly Glu Ser Gly Ala Glu Ala His  
 230 235 240  
 245 His His Asn His Phe Arg Gly Leu Gly Leu His Leu Gly Asp Ala Gly  
 250 255  
 260 Asp Gly Gly Arg Glu Ala Arg Val Ser Arg Tyr Arg Glu Lys Arg Arg  
 265 270  
 275 Thr Arg Leu Phe Ser Lys Lys Ile Arg Tyr Glu Val Arg Lys Leu Asn  
 280 285  
 290 Ala Glu Lys Arg Pro Arg Met Lys Lys Gly Arg Phe Val Lys Arg Ser Ser  
 295 300  
 305 Ile Gly Val Ala His

&lt;210&gt; 415

&lt;211&gt; 1300

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..1300:Ceres Seq. ID 2029806

&lt;400&gt;415

ctcaagttc ttgcaattac ctactacac caaagcaact cgaactgaac aaacacatcc  
 atggctctct ctggcaacc tctctctg agtcttcgg atgaaaata tgcgaatgc  
 agtgggaag agtggcatt caagtctgt cgtacgatt atgtatgt tgcgaatgc  
 aatcatggag agagtttca agagggaag attctctt ttgcattt gcaactaac  
 ctctgcctg ctgttttca gtagggcag gttttatg aaggtacaa agcttaacg  
 acagaagatg gtccgattct gctattcca ccagacaaa acgtctccg ccttaacgc  
 ggaagctgca gactctat gctctatct tctgtgatc aatgtctc cgcctacaa  
 caagtgtctc ttgcaacaa gaaatggat cctctccgg ggaaggaac attgtatatt  
 aggcctatct tgttgggag tggccgatt ctgtgttcat ttccattcc tgaacaccc  
 ttcaacgctt ttgcctgcc tgttggaact tatcataagg ataactctg ttgaattg  
 aaaaatcgaag atcagtttc tgcagtttt cctagtgaa ctgtgtgtg gaagagatc  
 acaaacattat gtcctgttg gataccattg gcagaggcga aaaaacaag ttctctgat  
 atttgtttt tggatgtgc aactggcaa acaattgaag aactttcgc agctaattt  
 ttatgtctca agggcaatg tgtatgaca ccaacaattg caggaaactat ttgtccgga  
 gtcaactgaa actgcgtaat ggaattgtgt cgtgatttcg gctaccaggt cgaggaaact

acgattcttc tagtgactt tctgatgcy gacgaagctt tctgactggs cactgottcc 960  
 atttgacta gtattgcatc cgtacacttt aaagacaaaa agaccggatt caaacaggg 1020  
 gaagaaact tggctcgaa gctatacag agttaagtg atatacagac ggtcgggtc 1080  
 gagataacca agggatggac gggagagatt gaccgcagg gctgaagtt gaaactgtaa 1140  
 cttgatgcta aatagtgtg tgtgtgtata caaaactat gtaagaaca tctgaagtg 1200  
 tcttgatct tctgatgtg gatgatcag catgctcat cgtgctgatg tattttataa 1260  
 atgttgctt gtaagtatt taataactat ggccttttgc 1300

<210> 416  
 <211> 374  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..374:Ceres Seq. ID 2029807

<400>416  
 Leu Lys Phe Leu Ala Ile Thr Tyr Leu Gln Ser Asn Ser Thr Arg 15  
 1 Thr Asn Thr Ser Met Ala Pro Ser Ala Gln Pro Leu Pro Val Ser Val 30  
 20 Ser Asp Glu Lys Tyr Ala Asn Val Lys Trp Glu Glu Leu Ala Phe Lys 45  
 35 Phe Val Arg Thr Asp Tyr Met Tyr Val Ala Lys Cys Asn His Gly Glu 60  
 50 Ser Phe Gln Glu Gly Lys Ile Leu Pro Phe Ala Asp Leu Leu Asn 75  
 65 Pro Cys Ala Ala Val Leu Gln Tyr Gly Gln Gly Leu Tyr Glu Gly Leu 90  
 85 Lys Ala Tyr Arg Thr Glu Asp Gly Arg Ile Leu Leu Phe Arg Pro Asp 110  
 100 Gln Asn Gly Leu Arg Leu Gln Ala Gly Ala Asp Arg Leu Tyr Met Pro 125  
 115 Tyr Pro Ser Val Asp Gln Phe Val Ser Ala Ile Lys Lys Gln Val Ala Leu 140  
 130 Ala Asn Lys Lys Trp Ile Pro Pro Gly Lys Gly Thr Leu Tyr Ile 155  
 145 Arg Pro Ile Leu Phe Gly Ser Gly Pro Ile Leu Gly Ser Phe Pro Ile 170  
 165 Pro Glu Thr Thr Phe Thr Ala Phe Ala Cys Pro Val Gly Arg Tyr His 185  
 195 Lys Asp Asn Ser Gly Leu Asn Leu Lys Ile Glu Asp Gln Phe Arg Arg 205  
 210 Ala Phe Pro Ser Gly Thr Gly Gly Val Lys Ser Ile Thr Asn Tyr Cys 220  
 225 Pro Val Trp Ile Pro Leu Ala Glu Ala Lys Lys Lys Gln Gly Phe Ser Asp 235  
 240 Ile Leu Phe Leu Asp Ala Ala Thr Gly Lys Asn Ile Glu Glu Leu Phe 255  
 245 Ala Ala Asn Val Phe Met Leu Lys Gly Asn Val Val Ser Thr Pro Thr 270  
 265 Ile Ala Gly Thr Ile Leu Pro Gly Val Thr Arg Asn Cys Val Met Glu 285  
 275 Leu Cys Arg Asp Phe Gly Tyr Gln Val Glu Glu Arg Thr Ile Pro Leu 300  
 290 Val Asp Phe Leu Asp Ala Asp Glu Ala Phe Cys Thr Gly Thr Ala Ser 315  
 305 Ile Val Thr Ser Ile Ala Ser Val Thr Phe Lys Asp Lys Lys Thr Gly 330  
 325

Phe Lys Thr Gly Glu Glu Thr Leu Ala Ala Lys Leu Tyr Glu Thr Leu 340  
 345 Ser Asp Ile Gln Thr Gly Arg Val Glu Asp Thr Lys Gly Trp Thr Val 360  
 355 Glu Ile Asp Arg Gln Gly 370  
 <210> 417  
 <211> 354  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..354:Ceres Seq. ID 2029808  
 <400>417  
 Met Ala Pro Ser Ala Gln Pro Leu Pro Val Ser Val Ser Asp Glu Lys 15  
 1 Tyr Ala Asn Val Lys Trp Glu Glu Leu Ala Phe Lys Phe Val Arg Thr 25  
 20 Asp Tyr Met Tyr Val Ala Lys Cys Asn His Gly Glu Ser Phe Gln Glu 40  
 35 Gly Lys Ile Leu Pro Phe Ala Asp Leu Gln Leu Asn Pro Cys Ala Ala 60  
 50 Val Leu Gln Tyr Gly Gln Gly Leu Tyr Glu Gly Leu Lys Ala Tyr Arg 75  
 65 Thr Glu Asp Gly Arg Ile Leu Leu Phe Arg Pro Asp Gln Asn Gly Leu 90  
 85 Arg Leu Gln Ala Gly Ala Asp Arg Leu Tyr Met Pro Tyr Pro Ser Val 110  
 100 Asp Gln Phe Val Ser Ala Ile Lys Gln Val Ala Leu Ala Asn Lys Lys 125  
 115 Trp Ile Pro Pro Gly Lys Gly Thr Leu Tyr Ile Arg Pro Ile Leu 140  
 130 Phe Gly Ser Gly Pro Ile Leu Gly Ser Phe Pro Ile Pro Glu Thr Thr 155  
 145 Phe Thr Ala Phe Ala Cys Pro Val Gly Arg Tyr His Lys Asp Asn Ser 170  
 165 Gly Leu Asn Leu Lys Ile Glu Asp Gln Phe Arg Arg Ala Phe Pro Ser 185  
 180 Gly Thr Gly Gly Val Lys Ser Ile Thr Asn Tyr Cys Pro Val Trp Ile 200  
 195 Pro Leu Ala Glu Ala Lys Lys Gln Gly Phe Ser Asp Ile Leu Phe Leu 220  
 210 Asp Ala Ala Thr Gly Lys Asn Ile Glu Glu Leu Phe Ala Ala Asn Val 235  
 225 Phe Met Leu Lys Gly Asn Val Val Ser Thr Pro Thr Ile Ala Gly Thr 250  
 245 Ile Leu Pro Gly Val Thr Arg Asn Cys Val Met Glu Leu Cys Arg Asp 265  
 260 Phe Gly Tyr Gln Val Glu Glu Arg Thr Ile Pro Leu Val Asp Phe Leu 280  
 275 Asp Ala Asp Glu Ala Phe Cys Thr Gly Thr Ala Ser Ile Val Thr Ser 300  
 290 Ile Ala Ser Val Thr Phe Lys Asp Lys Lys Thr Gly Phe Lys Thr Gly 315  
 305 Glu Glu Thr Leu Ala Ala Lys Leu Tyr Glu Thr Leu Ser Asp Ile Gln 330  
 325 Thr Gly Arg Val Glu Asp Thr Lys Gly Trp Thr Val Glu Ile Asp Arg 335



Gln Gly 340 345 350

<210> 418

<211> 320

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..320:Ceres Seq. ID 2029809

<400>418

Met Tyr Val Ala Lys Cys Asn His Gly Glu Ser Phe Gln Glu Gly Lys

1 15

Ile Leu Pro Phe Ala Asp Leu Gln Asn Pro Cys Ala Ala Val Leu

20 30

Gln Tyr Gly Gln Gly Leu Tyr Glu Gly Leu Lys Ala Tyr Arg Thr Glu

35 45

Asp Gly Arg Ile Leu Leu Phe Arg Pro Asp Gln Asn Gly Leu Arg Leu

50 60

Gln Ala Gly Ala Asp Arg Leu Tyr Met Pro Tyr Pro Ser Val Asp Gln

65 80

Phe Val Ser Ala Ile Lys Gln Val Ala Leu Ala Asn Lys Lys Trp Ile

85 95

Pro Pro Gly Lys Gly Thr Leu Tyr Ile Arg Pro Ile Leu Phe Gly

100 110

Ser Gly Pro Ile Leu Gly Ser Phe Pro Ile Pro Glu Thr Thr Phe Thr

115 125

Ala Phe Ala Cys Pro Val Gly Arg Tyr His Lys Asp Asn Ser Gly Leu

130 140

Asn Leu Lys Ile Glu Asp Gln Phe Arg Arg Ala Phe Pro Ser Gly Thr

145 155

Gly Gly Val Lys Ser Ile Thr Asn Tyr Cys Pro Val Trp Ile Pro Leu

165 175

Ala Glu Ala Lys Lys Gln Gly Phe Ser Asp Ile Leu Phe Leu Asp Ala

180 190

Ala Thr Gly Lys Asn Ile Glu Glu Leu Phe Ala Ala Asn Val Phe Met

195 205

Leu Lys Gly Asn Val Val Ser Thr Pro Thr Ile Ala Gly Thr Ile Leu

210 220

Pro Gly Val Thr Arg Asn Cys Val Met Glu Leu Cys Arg Asp Phe Gly

225 235

Tyr Gln Val Glu Glu Arg Thr Ile Pro Leu Val Asp Phe Leu Asp Ala

245 255

Asp Glu Ala Phe Cys Thr Gly Thr Ala Ser Ile Val Thr Ser Ile Ala

260 270

Ser Val Thr Phe Lys Asp Lys Lys Thr Gly Phe Lys Thr Gly Glu Glu

275 285

Thr Leu Ala Ala Lys Leu Tyr Glu Thr Leu Ser Asp Ile Gln Thr Gly

290 300

Arg Val Glu Asp Thr Lys Gly Trp Thr Val Glu Ile Asp Arg Gln Gly

305 315

<210> 419

<211> 987

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..987:Ceres Seq. ID 2030038

<400>419

atatctgaaa gactcaaaaa cctatctgca ttatctctca acaggcgga taacggagat

ggcgccatt acagctctca ctctcgctc tctctgttat cctccatcat ctgcacaaag

ccctagatc catggtctca ccaatcaacc accacagctt cgtctctctt ttctctctaa

cccttccct tctctatcaa tcaaaaaacc taatccatc cgaatttcg catcogcttc

ggcgtaaca acacaaatcc tcaagtcgga gaagtcacaa gctcgtctat cgaactcaac

cgggtccact cggctctctg ctactctgc ggccttagca atcgctgttaa ccaggtccct

agctcagaaa ctctctctcg caatccagac tccaagtcgc gtaatcgccg acgggttaacg

attctctctc agtaccgcgg gacctgtctt ctctcgctct ctccgggac gtcccctggg

atacttgaac acgcgcctga cgggtgttgc ggtggggata aagaagtggc tagacattta

cagtggggta ttgagtgtta gggttttgct agtttgttct cctaataacc ttgggaaag

acagctcttg tctgcaatla gagatctctg tgatccctac ttgaatctct tcagaaacat

tactctctct atcttcgata cgtttgtgtg tagtccatg ctgtctctcg cggttcttgg

tactcttga aacttctcat tgggaattt ctgcaactg ttcccaaga ttagggaata

atcagaattt gatctgtgtt tggatctctg atgagagtg tgaagtttat cgttttgtt

gtctctcttt tgactttgat tctgaataat gagatttgg gttgtgttt attctccaat

tacatgttca gactaccgtt tattgat

<210> 420

<211> 251

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..251:Ceres Seq. ID 2030039

<400>420

Tyr Leu Lys Asp Ser Lys Thr Tyr Arg His Leu Ser Ser Thr Gly Gly

1 5 10 15

Ile Thr Glu Met Ala Ala Ile Thr Ala Leu Thr Leu Arg Ser Pro Val

20 25 30

Tyr Leu Pro Ser Ser Ala Thr Ser Pro Arg Phe His Gly Phe Thr Asn

35 40 45

Gln Pro Pro Pro Ala Arg Leu Phe Phe Pro Leu Asn Pro Phe Pro Ser

50 55 60

Leu Ser Ile Gln Asn Pro Lys Ser Ile Arg Ile Ser Ala Ser Ala Ser

65 70 75 80

Pro Ile Thr Thr Pro Ile Leu Gln Thr Glu Lys Ser Thr Ala Arg Ser

85 90 95

Ser Thr Leu Thr Gly Ser Thr Arg Ser Leu Ala Thr Leu Ala Ala Leu

100 105 110

Ala Ile Ala Val Thr Arg Val Leu Ala Gln Lys Leu Ser Leu Ala Ile

115 120 125

Gln Thr Ser Ser Pro Val Ile Ala Asp Gly Leu Arg Phe Ser Leu Ser

130 135 140

Thr Ala Gly Pro Val Phe Phe Ala Ser Leu Arg Arg Pro Pro Gly

145 150 155 160

Tyr Leu Asn Thr Pro Leu Thr Val Val Ala Val Gly Ile Lys Lys Trp

165 170 175

Leu Asp Ile Tyr Ser Gly Val Leu Met Val Arg Val Leu Leu Ser Trp

180 185 190

Phe Pro Asn Ile Pro Trp Glu Arg Gln Pro Leu Ser Ala Ile Arg Asp

195 200 205

Leu Cys Asp Pro Tyr Leu Asn Leu Phe Arg Asn Ile Ile Pro Pro Ile  
210 215 220  
Phe Asp Thr Leu Asp Val Ser Pro Leu Leu Ala Phe Ala Val Leu Gly  
225 230 235 240  
Thr Leu Gly Ser Ile Val His Gly Ser Thr Gly  
245 250

<210> 421

<211> 232

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..232:Ceres Seq. ID 2030040

<400>421

Met Ala Ala Ile Thr Ala Leu Thr Leu Arg Ser Pro Val Tyr Leu Pro  
1 5 10 15

Ser Ser Ala Thr Ser Pro Arg Phe His Gly Phe Thr Asn Gln Pro Pro  
20 25 30

Pro Ala Arg Leu Phe Phe Pro Leu Asn Pro Phe Pro Ser Leu Ser Ile  
35 40 45

Gln Asn Pro Lys Ser Ile Arg Ile Ser Ala Ser Ala Ser Pro Ile Thr  
50 55 60

Thr Pro Ile Leu Gln Thr Glu Lys Ser Thr Ala Arg Ser Ser Thr Leu  
65 70 75 80

Thr Gly Ser Thr Arg Ser Leu Ala Thr Leu Ala Ala Leu Ala Ile Ala  
85 90 95

Val Thr Arg Val Leu Ala Gln Lys Leu Ser Leu Ala Ile Gln Thr Ser  
100 105 110

Ser Pro Val Ile Ala Asp Gly Leu Arg Phe Ser Leu Ser Thr Ala Gly  
115 120 125

Pro Val Phe Phe Ala Ser Leu Arg Asp Arg Pro Pro Gly Tyr Leu Asn  
130 135 140

Thr Pro Leu Thr Val Val Ala Val Gly Ile Lys Lys Trp Leu Asp Ile  
145 150 155 160

Tyr Ser Gly Val Leu Met Val Arg Val Leu Leu Ser Trp Phe Pro Asn  
165 170 175

Ile Pro Trp Glu Arg Gln Pro Leu Ser Ala Ile Arg Asp Leu Cys Asp  
180 185 190

Pro Tyr Leu Asn Leu Phe Arg Asn Ile Ile Pro Pro Ile Phe Asp Thr  
195 200 205

Leu Asp Val Ser Pro Leu Leu Ala Phe Ala Val Leu Gly Thr Leu Gly  
210 215 220

Ser Ile Val His Gly Ser Thr Gly  
225 230

<210> 422

<211> 755

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..755:Ceres Seq. ID 2031778

<400>422

cgcgagaaa cccatttgct cccttttgct caagctcagct gcccttttgc agcgaaagct  
60

caaccacaaa gagcatgatt atctcagaga ecaatccqag agagatctgc aataaccttt  
120  
tcaagaggg agtatgcttt gtaagaagg attcaatctt cccaagcat ccgttgattg  
180  
atctacaaa ctgcaagtg attaagctca tgcagattt caagccaag gagtagctta  
240  
ggagacatt tgcgtgattg cattattatt ggtttcgac taatgaagga atgagttct  
300  
tgagaactta tttaacctt ccttcagatc ttgtccctgc tactttgaag aagctcgcta  
360  
agccggtgg tcttacctt ggtggccac ctggtgatcg ccaaaagagga ccaactcgct  
420  
ctgatgaga cgtgccaga ttgtgtgacc gtgatgata ccgtggagga ccagtggtg  
480  
gtgatgaaa ggtgtgagct ccagctgatt tccagccgtc ttccaagga gttgttggtg  
540  
ggcttggttt tggcgtggt gcagtcggtt acagtcgagc agcaccaatc gtttcaggg  
600  
tcccttgaaa aatttgctg catattgcga cgtatgaagg acagttttgt tttttgtct  
660  
agttctggtt gtgtaatga aatctggaat ctataacta tctattacct tcaacttggt  
720  
ttaacgaac aaatcctgta gtttcaggat tagtg  
755

<210> 423

<211> 201

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..201:Ceres Seq. ID 2031779

<400>423

Ala Glu Asn Pro Ile Ala Pro Leu Cys Ser Ser Ser Val Pro Leu Cys  
1 5 10 15

Ser Glu Ser Ser Thr Asn Lys Thr Met Ile Ile Ser Glu Asn Asn Arg  
20 25 30

Arg Glu Ile Cys Lys Tyr Leu Phe Lys Glu Gly Val Cys Phe Ala Lys  
35 40 45

Lys Asp Phe Asn Leu Pro Lys His Pro Leu Ile Asp Val Pro Asn Leu  
50 55 60

Gln Val Ile Lys Leu Met Gln Ser Phe Lys Ser Lys Glu Tyr Val Arg  
65 70 75 80

Glu Thr Phe Ala Trp Met His Tyr Tyr Trp Phe Leu Thr Asn Glu Gly  
85 90 95

Ile Glu Phe Leu Arg Thr Tyr Leu Asn Leu Pro Ser Asp Val Val Pro  
100 105 110

Ala Thr Leu Lys Lys Ser Ala Lys Pro Gly Arg Pro Phe Gly Gly  
115 120 125

Pro Pro Gly Asp Arg Gln Arg Gly Pro Pro Arg Ser Asp Gly Asp Arg  
130 135 140

Pro Arg Phe Gly Asp Arg Asp Gly Tyr Arg Gly Gly Pro Arg Gly Gly  
145 150 155 160

Asp Glu Lys Gly Gly Ala Pro Ala Asp Phe Gln Pro Ser Phe Gln Gly  
165 170 175

Gly Gly Gly Arg Pro Gly Phe Gly Arg Gly Ala Gly Tyr Ser Ala  
180 185 190

Ala Ala Pro Ser Gly Ser Gly Phe Pro  
195 200

<210> 424

<211> 177

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..177:Ceres Seq. ID 2031780

<400>424  
Met Ile Ile Ser Glu Asn Asn Arg Arg Glu Ile Cys Lys Tyr Leu Phe  
1 5 10 15  
Lys Glu Gly Val Cys Phe Ala Lys Lys Asp Phe Asn Leu Pro Lys His  
20 25 30  
Pro Leu Ile Asp Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln Ser  
35 40 45  
Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His Tyr  
50 55 60  
Tyr Trp Phe Leu Thr Asn Glu Gly Ile Glu Phe Leu Arg Thr Tyr Leu  
65 70 75 80  
Asn Leu Pro Ser Asp Val Val Pro Ala Thr Leu Lys Lys Ser Ala Lys  
85 90 95  
Pro Gly Gly Arg Pro Phe Gly Gly Pro Pro Gly Asp Arg Gln Arg Gly  
100 105 110  
Pro Pro Arg Ser Asp Gly Asp Arg Pro Arg Phe Gly Asp Arg Asp Gly  
115 120 125  
Tyr Arg Gly Gly Pro Arg Gly Gly Asp Glu Lys Gly Gly Ala Pro Ala  
130 135 140  
Asp Phe Gln Pro Ser Phe Gln Gly Gly Gly Arg Pro Gly Phe Gly  
145 150 155 160  
Arg Gly Ala Gly Tyr Ser Ala Ala Ala Pro Ser Gly Ser Gly Phe  
165 170 175  
Pro

<210> 425

<211> 332

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..132:Ceres Seq. ID 2031781

<400>425  
Met Gln Ser Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp  
1 5 10 15  
Met His Tyr Tyr Trp Phe Leu Thr Asn Glu Gly Ile Glu Phe Leu Arg  
20 25 30  
Thr Tyr Leu Asn Leu Pro Ser Asp Val Val Pro Ala Thr Leu Lys Lys  
35 40 45  
Ser Ala Lys Pro Gly Gly Arg-Pro Phe Gly Gly Pro Gly Asp Arg  
50 55 60  
Gln Arg Gly Pro Pro Arg Ser Asp Gly Asp Arg Pro Arg Phe Gly Asp  
65 70 75 80  
Arg Asp Gly Tyr Arg Gly Gly Pro Arg Gly Gly Asp Glu Lys Gly Gly  
85 90 95  
Ala Pro Ala Asp Phe Gln Pro Ser Phe Gln Gly Gly Gly Arg Pro  
100 105 110  
Gly Phe Gly Arg Gly Ala Gly Gly Tyr Ser Ala Ala Pro Ser Gly  
115 120 125  
Ser Gly Phe Pro  
130

<210> 426

<211> 538

<212> DNA

<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown  
  
<223> LOCATION: 1..538:Ceres Seq. ID 2032723  
  
<400>426  
aaagaataa aaaaacaaa tggcgtttac tccgaagatc atcacatgcc tcatgtcct  
tacgatctac atggcatccc caacagatgc aaccatccag tgtgggacag tgacgagcac  
atggacacag tgcctgacct aattgaccaa cagtgtcca ttgccatcac aatgtgcgt  
ggagctaacg tcatgttacc aattgctcca gaccacacag gaccataaac aagtatgga  
gtgcctaaa atgaggaata agaaatcaa gggcctcaac accgaccttg tggcgcact  
tctacacact tgtgtgtt caattcccta ccccatcagt tttagcacca atgcgcacg  
tctatcact gcgctgaa agaggttagt gatcagatgt acgactaatc aaacttgcca  
gctttaacc kaattaata aaagtattct gctataatt cccatttat gattttatc  
tctatctat gtaaccacac gatttcatat gctaataatg acaacggatc ttctctc  
60  
120  
180  
240  
300  
360  
420  
480  
538

<210> 427

<211> 119

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..119:Ceres Seq. ID 2032724

<400>427  
Met Ala Phe Thr Pro Lys Ile Ile Thr Cys Leu Ile Val Leu Thr Ile  
1 5 10 15  
Tyr Met Ala Ser Pro Thr Glu Ser Thr Ile Gln Cys Gly Thr Val Thr  
20 25 30  
Ser Thr Leu Ala Gln Cys Leu Thr Tyr Leu Thr Asn Ser Gly Pro Leu  
35 40 45  
Pro Ser Gln Cys Cys Val Gly Val Lys Ser Leu Tyr Gln Leu Ala Gln  
50 55 60  
Thr Thr Pro Asp Arg Lys Gln Val Cys Glu Cys Leu Lys Leu Ala Gly  
65 70 75 80  
Lys Glu Ile Lys Gly Leu Asn Thr Asp Leu Val Ala Ala Leu Pro Thr  
85 90 95  
Thr Cys Gly Val Ser Ile Pro Tyr Pro Ile Ser Phe Ser Thr Asn Cys  
100 105 110  
Asp Ser Ile Ser Thr Ala Val  
115

<210> 428

<211> 102

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..102:Ceres Seq. ID 2032725

<400>428  
Met Ala Ser Pro Thr Glu Ser Thr Thr Ile Gln Cys Gly Thr Val Thr Ser  
1 5 10 15  
Thr Leu Ala Gln Cys Leu Thr Tyr Leu Thr Asn Ser Gly Pro Leu Pro  
20 25 30  
Ser Gln Cys Cys Val Gly Val Lys Ser Leu Tyr Gln Leu Ala Gln Thr  
35 40 45

Thr Pro Asp Arg Lys Gln Val Cys Glu Cys Leu Lys Leu Ala Gly Lys  
50 55 60  
Glu Ile Lys Gly Leu Asn Thr Asp Leu Val Ala Leu Pro Thr Thr  
65 70 75 80  
Cys Gly Val Ser Ile Pro Tyr Pro Ile Ser Phe Ser Thr Asn Cys Asp  
85 90 95  
Ser Ile Ser Thr Ala Val  
100

&lt;210&gt; 429

&lt;211&gt; 75

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..75:Ceres Seq. ID 2032726

<400>429  
Lys Ala Ile Lys Asn Lys Asn Gly Val Tyr Ser Glu Asp His His Met  
1 5 10 15  
Pro His Cys Pro Tyr Asp Leu His Gly Ile Pro Asn Arg Val Asn His  
20 25 30  
Pro Val Trp Asp Ser Asp Glu His Thr Gly Thr Val Pro Asp Leu  
35 40 45  
Asp Gln Gln Trp Ser Ile Ala Ile Thr Met Leu Arg Gly Ser Gln Val  
50 55 60  
Ile Val Pro Ile Gly Ser Asp His Thr Gly Pro  
65 70 75

&lt;210&gt; 430

&lt;211&gt; 198

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..198:Ceres Seq. ID 2035536

<400>430  
atgctgttg tatgttgga agcagatgttg cctctcgga teatcgtgg gatgctctgt  
60  
atcattggca attctcagta ctacatccac aaagcttatc atggcgtcc taagcaatc  
120  
ggcacgagt aatgggatgt tgcattggaa agacgcgaca agaagtgct cgagaagct  
180  
gcagctcctt cctcatga  
198

&lt;210&gt; 431

&lt;211&gt; 65

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..65:Ceres Seq. ID 2035537

<400>431  
Met Ser Leu Val Trp Leu Glu Ala Met Leu Pro Leu Gly Ile Ile Gly  
1 5 10 15  
Gly Met Leu Cys Ile Met Gly Asn Ser Gln Tyr Thr Ile His Lys Ala

Tyr His Gly Arg Pro Lys His Ile Gly His Asp Glu Trp Asp Val Ala  
20 25 30 35 40 45  
Met Glu Arg Arg Asp Lys Lys Val Val Glu Lys Ala Ala Pro Ser  
50 55 60  
Ser  
65

&lt;210&gt; 432

&lt;211&gt; 65

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..65:Ceres Seq. ID 2035538

<400>432  
Cys Arg Trp Tyr Gly Trp Lys Arg Cys Cys Leu Ser Glu Ser Ser Val  
1 5 10 15  
Gly Cys Ser Val Ser Trp Ala Ile Leu Ser Thr Thr Ser Thr Lys Leu  
20 25 30  
Ile Met Ala Val Leu Ser Thr Ser Ala Thr Met Asn Gly Met Leu  
35 40 45  
Trp Lys Asp Ala Thr Arg Lys Ser Ser Arg Lys Leu Gln Leu Leu Pro  
50 55 60  
His  
65

&lt;210&gt; 433

&lt;211&gt; 57

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..57:Ceres Seq. ID 2035539

<400>433  
Met Leu Pro Leu Gly Ile Ile Gly Met Leu Cys Ile Met Gly Asn  
1 5 10 15  
Ser Gln Tyr Tyr Ile His Lys Ala Tyr His Gly Arg Pro Lys His Ile  
20 25 30  
Gly His Asp Glu Trp Asp Val Ala Met Glu Arg Arg Asp Lys Lys Val  
35 40 45  
Val Glu Lys Ala Ala Pro Ser Ser  
50 55

&lt;210&gt; 434

&lt;211&gt; 573

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..573:Ceres Seq. ID 2035575

&lt;400&gt;434

atgagagatt ctatgagatg taccagcaat tccaagtcta tgaagattca tgcacaaagag  
 aaggttccgg tgaactcaaa aaccattatg cagattcatg gagagttaga tactggaact  
 ggagcccgga gtaactctg tgcgatgatt agacactttt ttctgagcg ttcaacagcg  
 ctgggggttg gatgattata tgaatagcgc caaagcttc ggtgtcttgc acgcggaaaa  
 aagagcttc ctgaagagc tgaatagcgt gtaaccttta atattaaagg gcggtgtgat  
 atgatcagg actaaatca gaagaacccc aaaggagcag cagaattgc ctgaacata  
 atgatttca aggaagatca ggatgcagg atcaaatgt gctacgaat gtttgataag  
 gtccctata tgcagattag agaaacaaat tgaactctca acgcgaacat gaagggaata  
 tgaacttgc ggaatgacct gaaactgat tttttcaat catcatctga gaaatgtatt  
 gataccactg ctgatgaaca catttcaatt cta

<210> 435

<211> 167

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..167:Ceres Seq. ID 2035576

<400>435

Met Glu Thr Ser Met Arg Tyr Thr Ser Asn Ser Lys Ser Met Lys Ile  
 1 5 10 15  
 His Ala Lys Glu Lys Val Pro Val Asn Ser Lys Thr His Leu Gln Leu  
 20 25 30  
 His Gly Glu Leu Asp Thr Gly Thr Gly Ala Pro Ser Tyr Thr Phe Cys Ala  
 35 40 45  
 Met Ile Arg His Phe Phe Pro Glu Ala Ser Thr Gly Leu Gly Val Gly  
 50 55 60  
 Leu His Tyr Asp Lys Arg Gln Lys Leu Arg Cys Leu Val Arg Gly Lys  
 65 70 75  
 Lys Glu Phe Pro Val Arg Ala Asp Lys Arg Val Thr Phe Asn Ile Lys  
 80 85 90  
 Gly Arg Cys Asp Ile Asp Gln Asp Leu Asn Gln Lys Asn Pro Lys Gly  
 95 100 105  
 Ala Ala Glu Phe Ala Trp Asn Ile Met Asp Phe Lys Glu Asp Gln Asp  
 110 115 120  
 Val Arg Ile Lys Val Gly Tyr Glu Met Phe Asp Lys Val Pro Tyr Met  
 125 130 135  
 Gln Ile Arg Glu Asn Asn Trp Thr Leu Asn Ala Asn Met Lys Gly Lys  
 140 145 150  
 Trp Asn Leu Arg Tyr Asp Leu  
 155 160 165

<210> 436

<211> 163

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..163:Ceres Seq. ID 2035577

<400>436

Met Arg Tyr Thr Ser Asn Ser Lys Ser Met Lys Ile His Ala Lys Glu  
 1 5 10  
 Lys Val Pro Val Asn Ser Lys Thr His Leu Gln Leu His Gly Glu Leu  
 15 20 25 30  
 Asp Thr Gly Thr Gly Ala Pro Ser Tyr Phe Cys Ala Met Ile Arg His

35 40 45  
 Phe Phe Pro Glu Ala Ser Thr Gly Leu Gly Val Gly Leu His Tyr Asp  
 50 55 60  
 Lys Arg Gln Lys Leu Arg Cys Leu Val Arg Gly Lys Lys Glu Phe Pro  
 65 70 75 80  
 Val Arg Ala Asp Lys Arg Val Thr Phe Asn Ile Lys Gly Arg Cys Asp  
 85 90 95  
 Ile Asp Gln Asp Leu Asn Gln Lys Asn Pro Lys Gly Ala Ala Glu Phe  
 100 105 110  
 Ala Trp Asn Ile Met Asp Phe Lys Glu Asp Gln Asp Val Arg Ile Lys  
 115 120 125  
 Val Gly Tyr Glu Met Phe Asp Lys Val Pro Tyr Met Gln Ile Arg Glu  
 130 135 140  
 Asn Asn Trp Thr Leu Asn Ala Asn Met Lys Gly Lys Trp Asn Leu Arg  
 145 150 155 160  
 Tyr Asp Leu

<210> 437

<211> 154

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..154:Ceres Seq. ID 2035578

<400>437  
 Met Lys Ile His Ala Lys Glu Lys Val Pro Val Asn Ser Lys Thr His  
 1 5 10 15  
 Leu Gln Leu His Gly Glu Leu Asp Thr Gly Thr Gly Ala Pro Ser Tyr  
 20 25 30  
 Phe Cys Ala Met Ile Arg His Phe Phe Pro Glu Ala Sor Thr Gly Leu  
 35 40 45  
 Gly Val Gly Leu His Tyr Asp Lys Arg Gln Lys Leu Arg Cys Leu Val  
 50 55 60  
 Arg Gly Lys Lys Glu Phe Pro Val Arg Ala Asp Lys Arg Val Thr Phe  
 65 70 75  
 Asn Ile Lys Gly Arg Cys Asp Ile Asp Gln Asp Leu Asn Gln Lys Asn  
 80 85 90  
 Pro Lys Gly Ala Ala Glu Phe Ala Trp Asn Ile Met Asp Phe Lys Glu  
 95 100 105 110  
 Asp Gln Asp Val Arg Ile Lys Val Gly Tyr Glu Met Phe Asp Lys Val  
 115 120 125  
 Pro Tyr Met Gln Ile Arg Glu Asn Asn Trp Thr Leu Asn Ala Asn Met  
 130 135 140  
 Lys Gly Lys Trp Asn Leu Arg Tyr Asp Leu  
 145 150 155

<210> 438

<211> 750

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..750:Ceres Seq. ID 2036457

<400>438

atgacgacc agatcagcaa gaagagaag ttgtagcgg acggtgtgt ctacgtgaa 60  
 ttgaatggt ttctaacg agagtagc gagatggt actctgggt tgaggttagg 120  
 gttaactga ttgagacta gattatcac agagtacac gtactcaga ttctctcgt 180  
 gagaaggga ggagaattg ggaatgact tccctgttc agaagagat caagtcca 240  
 gtacaggg ttgagctta tgcagaaag gttacaaca gaggtctcg tgcctcgt 300  
 cagctgagt cttacgtta caagcttct ggtggtcgc cgttctgag gcttgat 360  
 ggtgttga gttgttat ggagatgga gctaaagat gcaggttat cgtgagtg 420  
 aagctctg ctgccagc caagctatg aadccaag atggtacat ggtgcatt 480  
 ggtcaacca taagaata catagactt gcagtgcg atgtttgct tagacaagt 540  
 gtgtggaa tcaagtgaa ggtatgct gattggacc ctaaggcat atcaggaca 600  
 aagacacat tgcctatgt tggatcat cattctcta aagaagaaga ggcattat 660  
 gcaactgct aggttctgc ccgctgct cctgtagcag atgcaccat cacagccta 720  
 gattaccctg cgatgatccc agtcgccta 750

&lt;210&gt; 439

&lt;211&gt; 249

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..249:Ceres Seq. ID 2036458

&lt;400&gt;439

Met Thr Thr Gln Ile Ser Lys Lys Arg Lys Phe Val Ala Asp Gly Val 15  
 1 Phe Tyr Ala Glu Leu Asn Glu Val Leu Thr Arg Glu Leu Ala Glu Asp 30  
 Gly Tyr Ser Gly Val Glu Val Arg Val Thr Pro Met Arg Thr Glu Ile 45  
 35 Ile Arg Ala Thr Arg Thr Gln Asn Val Leu Gly Glu Lys Gly Arg 60  
 50 Arg Ile Arg Glu Leu Thr Ser Leu Val Gln Lys Arg Phe Lys Phe Pro 75  
 65 Val Asp Ser Val Glu Leu Tyr Ala Glu Lys Val Asn Asn Arg Gly Leu 90  
 85 Cys Ala Ile Ala Gln Ala Glu Ser Leu Arg Tyr Lys Leu Leu Gly Gly 105  
 100 Leu Ala Val Arg Arg Ala Cys Tyr Gly Val Leu Arg Phe Val Met Glu 115  
 115 Ser Gly Ala Lys Gly Cys Glu Val Ile Val Ser Gly Lys Leu Arg Ala 125  
 130 Ala Arg Ala Lys Ser Met Lys Phe Lys Asp Gly Tyr Met Val Ser Ser 145  
 145 Gly Gln Pro Thr Lys Glu Tyr Ile Asp Ser Ala Val Arg His Val Leu 155  
 160 Leu Arg Gln Gly Val Leu Gly Ile Lys Val Lys Val Met Leu Asp Trp 175  
 180 Asp Pro Lys Gly Ile Ser Gly Pro Lys Thr Pro Leu Pro Asp Val Val 185  
 190 Ile Ile His Ser Pro Lys Glu Glu Ala Ile Tyr Ala Pro Ala Gln 205  
 210 Val Ala Ala Pro Ala Ala Leu Val Ala Asp Ala Pro Leu Thr Ala Val 220  
 225 Asp Tyr Pro Ala Met Ile Pro Val Ala 235  
 245

&lt;210&gt; 440

&lt;211&gt; 206

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..206:Ceres Seq. ID 2036459

<400>440  
 Met Arg Thr Glu Ile Ile Arg Ala Thr Arg Thr Gln Asn Val Leu 15  
 1 Gly Glu Lys Gly Arg Arg Ile Arg Glu Leu Thr Ser Leu Val Gln Lys 30  
 20 Arg Phe Lys Phe Pro Val Asp Ser Val Glu Leu Tyr Ala Glu Lys Val 35  
 40 Asn Asn Arg Gly Leu Cys Ala Ile Ala Gln Ala Glu Ser Leu Arg Tyr 55  
 50 Lys Leu Leu Gly Gly Leu Ala Val Arg Arg Ala Cys Tyr Gly Val Leu 60  
 65 Arg Phe Val Met Glu Ser Gly Ala Lys Gly Cys Glu Val Ile Val Ser 75  
 80 Gly Lys Leu Arg Ala Ala Arg Ala Lys Ser Met Lys Phe Lys Asp Gly 85  
 90 Tyr Met Val Ser Ser Gly Gln Pro Thr Lys Glu Tyr Ile Asp Ser Ala 105  
 110 Val Arg His Val Leu Leu Arg Gln Gly Val Leu Gly Ile Lys Val Lys 115  
 120 Val Met Leu Asp Trp Asp Pro Lys Gly Ile Ser Gly Pro Lys Thr Pro 135  
 140 Leu Pro Asp Val Val Ile Ile His Ser Pro Lys Glu Glu Ala Ile 145  
 150 Tyr Ala Pro Ala Gln Val Ala Ala Pro Ala Ala Leu Val Ala Asp Ala 155  
 160 Pro Leu Thr Ala Val Asp Tyr Pro Ala Met Ile Pro Val Ala 175  
 180 185 190 205

&lt;210&gt; 441

&lt;211&gt; 549

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..549:Ceres Seq. ID 2036585

<400>441  
 atggttcca agacggagct ttgtcgttc agtgacaga agattacc aggaagaga 60  
 atagtttta tccgactga ttctcaggt ttctgttct ttaactcaaa atgtaagagg 120  
 tactcata acaggtgaa gccatccag ctgtcatgga ctgcctgta cagaagacaa 180  
 caaagaagg atgcagaca agagctgtg asagaagga gactgccac caagaagcca 240  
 tactcaagt ccattgttg tgcctcttg gaagtaattc agaagaagag agctagaag 300  
 cctgaagttc gtgatgcag cagggaagt gctctggtg agatcaagga aagaatcaaa 360  
 aagaccaag atgaagaag ggttaagag gtggaattg ctctcaagca acagaagttc 420  
 aagctaat tcccaagc tgtgtgca tccaagggtc ctaaggtggt ggtgcaaac 480  
 gctgaagagc ttaagccat cttctctac tctgctct ttctgctagt agtacttt 540  
 agtagttga 549

&lt;210&gt; 442

&lt;211&gt; 182

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..182:Ceres Seq. ID 2036586

<400>442  
 Met Val Leu Lys Thr Glu Leu Cys Arg Phe Ser Gly Gln Lys Ile Tyr  
 1 5 10 15  
 Pro Gly Arg Gly Ile Arg Phe Ile Arg Ser Asp Ser Gln Val Phe Leu  
 20 25 30  
 Phe Leu Asn Ser Lys Cys Arg Tyr Phe His Asn Lys Lys Lys Pro  
 35 40 45  
 Ser Lys Leu Ala Trp Thr Ala Met Tyr Arg Lys Gln His Lys Lys Asp  
 50 55 60  
 Ala Ala Gln Glu Ala Val Lys Arg Arg Arg Ala Thr Lys Lys Pro  
 65 70 75 80  
 Tyr Ser Arg Ser Ile Val Gly Ala Thr Leu Glu Val Ile Gln Lys Lys  
 85 90  
 Arg Ala Glu Lys Pro Glu Val Arg Asp Ala Ala Arg Glu Ala Ala Leu  
 100 105 110  
 Arg Glu Ile Lys Glu Arg Ile Lys Lys Thr Lys Asp Glu Lys Lys Ala  
 115 120  
 Lys Lys Val Glu Phe Ala Ser Lys Lys Gln Lys Val Lys Ala Asn Phe  
 130 135 140  
 Pro Lys Ala Ala Ala Ser Lys Gly Pro Lys Val Val Val Ala Asn  
 145 150 155 160  
 Ala Glu Glu Lys Lys Ala Ile Phe Ser His Ser Ala Ser Phe Leu Leu  
 165 170 175  
 Val Ala Thr Phe Ser Ser  
 180

<210> 443

<211> 127

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..127:Ceres Seq. ID 2036587

<400>443  
 Met Tyr Arg Lys Gln His Lys Lys Asp Ala Ala Gln Glu Ala Val Lys  
 1 5 10 15  
 Arg Arg Arg Arg Ala Thr Lys Lys Lys Pro Tyr Ser Arg Ser Ile Val Gly  
 20 25 30  
 Ala Thr Leu Glu Val Ile Gln Lys Lys Arg Ala Glu Lys Pro Glu Val  
 35 40 45  
 Arg Asp Ala Ala Arg Glu Ala Ala Leu Arg Glu Ile Lys Glu Arg Ile  
 50 55 60  
 Lys Lys Thr Lys Asp Glu Lys Lys Ala Lys Lys Val Glu Phe Ala Ser  
 65 70 75 80  
 Lys Gln Gln Lys Val Lys Ala Asn Phe Pro Lys Ala Ala Ala Ser  
 85 90  
 Lys Gly Pro Lys Val Val Ala Asn Ala Glu Glu Leu Lys Ala Ile  
 100 105 110  
 Phe Ser His Ser Ala Ser Phe Leu Leu Val Ala Thr Phe Ser Ser  
 115 120 125

<210> 444

<211> 486

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..486:Ceres Seq. ID 2039554

<400>444  
 aggcgtctcg tcggcgctcc aatagctgtc ccggcgcttc agctgcagac acaattctc  
 tccaatccca tctctccccc attctccggg tctttctcca ccggaatac accagcaact  
 tctctctcg tagctatggc tcccagaaa aagttgaaca aatgatgc caagtggaag  
 aaacaatggt acggagctgg attgttttc gaaggagtg agcaataaa ctttgatgtt  
 tccaagaagc tgggaagcgg aagaagtgtg agcaacgttg agaatactgg cctgtgtcca  
 aaagaagaag gattggact cacattgtca tctcttgaga agcttaaatg cttctccaaa  
 gcagagacc ttgtctctct cagtctctt gagaacttag ctggaacatc gcttcgggtc  
 ttagctctgg ctgcataacc agctctcaag agtgcctatt gttagcctgg tgttgatccc  
 ggaatga

<210> 445

<211> 161

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..161:Ceres Seq. ID 2039555

<400>445  
 Met Ala Val Val Gly Ala Pro Ile Ser Ser Pro Ala Ala Gln Leu Gln  
 1 5 10 15  
 Thr Gln Phe Leu Ser Asn Pro Ile Leu Pro Arg Phe Arg Arg Ser Phe  
 20 25 30  
 Ser Thr Gly Lys Ser Pro Ala Thr Phe Ser Val Val Ala Met Ala Pro  
 35 40 45  
 Gln Lys Lys Val Asn Lys Tyr Asp Ala Lys Trp Lys Lys Gln Trp Tyr  
 50 55 60  
 Gly Ala Gly Leu Phe Phe Glu Gly Ser Glu Gln Ile Asn Val Asp Val  
 65 70 75 80  
 Phe Lys Lys Leu Glu Lys Arg Lys Val Leu Ser Asn Val Glu Lys Ser  
 85 90 95  
 Gly Leu Leu Ser Lys Ala Glu Gly Leu Gly Leu Thr Leu Ser Ser Leu  
 100 105 110  
 Glu Lys Leu Lys Val Phe Ser Lys Ala Glu Asp Leu Gly Leu Leu Ser  
 115 120 125  
 Leu Leu Glu Asn Leu Ala Gly Thr Ser Pro Ala Val Leu Ala Ser Ala  
 130 135 140  
 Ala Leu Pro Ala Leu Thr Ser Cys Tyr Cys Ser Arg Gly Val Asp Pro  
 145 150 155 160  
 Gly

<210> 446

<211> 116

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..116:Ceres Seq. ID 2039556

<400>446

Met Ala Pro Gln Lys Lys Val Asn Lys Tyr Asp Ala Lys Trp Lys Lys  
1 5 10 15  
Gln Trp Tyr Gly Ala Gly Leu Phe Phe Gln Gly Ser Glu Gln Ile Asn  
20 25 30  
Val Asp Val Phe Lys Lys Leu Glu Lys Arg Lys Val Leu Ser Asn Val  
35 40 45  
Glu Lys Ser Gly Leu Leu Ser Lys Ala Glu Gly Leu Gly Leu Thr Leu  
50 55 60  
Ser Ser Leu Glu Lys Leu Lys Val Phe Ser Lys Ala Glu Asp Leu Gly  
65 70 75 80  
Leu Leu Ser Leu Leu Glu Asn Leu Ala Gly Thr Ser Pro Ala Val Leu  
85 90 95  
Ala Ser Ala Ala Leu Pro Ala Leu Thr Ser Cys Tyr Cys Ser Arg Gly  
100 105 110  
Val Asp Pro Gly  
115

<210> 447

<211> 378

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..378:Ceres Seq. ID 2044283

<400>447

atgcgaatt tgaatgag attaccaatt agcttgagaa gcttcttgt ttcagcttct  
120  
tcatccaag gtccgcgc ggtgacgga gatctagcg ggggtgtag accgatgatt  
180  
gtgaatrac cgttgagaaa gatacagaa cgttgatgc gaaccagatc caacgatcag  
240  
aacaagta aagagcttat ggtatgatac cgtcaaatcg gttctcaagt tccgatgat  
300  
gtgatgaag ttatgtgaac ttatgtgg ttctggatg gtcacagata cgaagcgcat  
360  
cagaagctag ggtctcaac tatacgttc aaatccgta aaggaaacaa ggaacatta  
378  
agcatcatc ttgcctga

<210> 448

<211> 125

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..125:Ceres Seq. ID 2044284

<400>448

Met Ala Asn Leu Met Met Arg Leu Pro Ile Ser Leu Arg Ser Phe Ser  
1 5 10 15  
Val Ser Ala Ser Ser Ser Asn Gly Ser Pro Pro Val Ile Gly Gly Ser  
20 25 30  
Ser Gly Gly Val Gly Pro Met Ile Val Glu Leu Pro Leu Glu Lys Ile  
35 40 45  
Arg Arg Pro Leu Met Arg Thr Arg Ser Asn Asp Gln Asn Lys Val Lys  
50 55 60  
Glu Leu Met Asp Ser Ile Arg Gln Ile Gly Leu Gln Val Pro Ile Asp  
65 70 75 80

Val Ile Glu Val Asp Gly Thr Tyr Tyr Gly Phe Ser Gly Cys His Arg  
85 90 95  
Tyr Glu Ala His Gln Lys Leu Gly Leu Pro Thr Ile Arg Cys Lys Ile  
100 105 110  
Arg Lys Gly Thr Lys Glu Thr Leu Arg His His Leu Arg  
115 120 125

<210> 449

<211> 121

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..121:Ceres Seq. ID 2044285

<400>449

Met Met Arg Leu Pro Ile Ser Leu Arg Ser Phe Ser Val Ser Ala Ser  
1 5 10 15  
Ser Ser Asn Gly Ser Pro Pro Val Ile Gly Gly Ser Ser Gly Gly Val  
20 25 30  
Gly Pro Met Ile Val Glu Leu Pro Leu Glu Lys Ile Arg Arg Pro Leu  
35 40 45  
Met Arg Thr Arg Ser Asn Asp Gln Asn Lys Val Lys Glu Leu Met Asp  
50 55 60  
Ser Ile Arg Gln Ile Gly Leu Gln Val Pro Ile Asp Val Ile Glu Val  
65 70 75 80  
Asp Gly Thr Tyr Tyr Gly Phe Ser Gly Cys His Arg Tyr Glu Ala His  
85 90 95  
Gln Lys Leu Gly Leu Pro Thr Ile Arg Cys Lys Ile Arg Lys Gly Thr  
100 105 110  
Lys Glu Thr Leu Arg His His Leu Arg  
115 120

<210> 450

<211> 120

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..120:Ceres Seq. ID 2044286

<400>450

Met Arg Leu Pro Ile Ser Leu Arg Ser Phe Ser Val Ser Ala Ser Ser  
1 5 10 15  
Ser Asn Gly Ser Pro Pro Val Ile Gly Gly Ser Ser Gly Gly Val Gly  
20 25 30  
Pro Met Ile Val Glu Leu Pro Leu Glu Lys Ile Arg Arg Pro Leu Met  
35 40 45  
Arg Thr Arg Ser Asn Asp Gln Asn Lys Val Lys Glu Leu Met Asp Ser  
50 55 60  
Ile Arg Gln Ile Gly Leu Gln Val Pro Ile Asp Val Ile Glu Val Asp  
65 70 75 80  
Gly Thr Tyr Tyr Gly Phe Ser Gly Cys His Arg Tyr Glu Ala His Gln  
85 90 95  
Lys Leu Gly Leu Pro Thr Ile Arg Cys Lys Ile Arg Lys Gly Thr Lys  
100 105 110  
Glu Thr Leu Arg His His Leu Arg



115 120

&lt;210&gt; 451

&lt;211&gt; 1236

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..1236:Ceres Seq. ID 2048114

&lt;400&gt;451

atagaagac ctagtcaaat gatgggctt ctattacat ccttttcgg tgtattgtt  
 gtttcccta tgggtattac ttttcaacc ttgactttta ctgaatgaa tcttccatc  
 acatttttc cctcagatga tcttcatac attgagata aatactctga catatcaaga  
 caaagactat tgggtcttg gtcttcaca aaaggcccta aactcaagaa tgcattccct  
 gacctccat ataactataa tgaactaag gtgatgata gaacgttga gctattgcaag  
 atatgggtt cgaactaacc cgtgtgtgt gtagggtac caccagatat agtcacgct  
 gatcagatt ttactctcg tgcactgtg ggcgacctac atgaggattt aacagtcaag  
 cagggtatc tagtaacatt tacgttggc tatgatcaga ggaataatat agacacttg  
 ttgaagaagt tctcagataa ctctctata atgtgtttc actacgatg ccgggcaagc  
 gaatgggaag atttgaatt gtccaagcga gccattcatt tgagcatcog gaacaaca  
 aaatgggtt agcacaagc atttctcat cctgacatag ttgcccccta tgaatatc  
 ttcattatgg atgagatct tggcttgaa cactttgatt cggcaataa tctggcgtg  
 gtgaagaag atggtttga atcttcacg cctgattag agccataga agggctcaca  
 tggagatga caaagaanaa agacgaact gaatccaca agcatgcta ggaagaat  
 ggtgtgtca gcatcccaa ttaccctct tggcagcgt ttggtgagat tatggctct  
 gttttccc ggaagcatg ggcgtgtgt tggcatatga ttcaagaca ttgtattcat  
 ggaatggctc tggacttgc ctttcgaaa tggattcaga acgcacaga gaaatttga  
 gtgtagatg ctaatggat tatactcaa ggtgtccat cettatggaa tcaaggaca  
 ccagagaag ggaacaacc atgggaagg gtagagaac gatcgaggag agagtggaca  
 atgttcaa acagattga tgaactgaa aaagcttatt ttgaagcatc tgcacaag  
 aaagtctt caggcctca cgggaattg gtagat

&lt;210&gt; 452

&lt;211&gt; 411

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..411:Ceres Seq. ID 2048115

&lt;400&gt;452

Met Arg Arg Pro Ser Gln Met Met Arg Leu Leu Thr Ser Phe Phe  
 1 5 10 15  
 Gly Val Ile Val Gly Phe Leu Met Gly Ile Thr Phe Pro Thr Leu Thr  
 20 25 30  
 Leu Thr Lys Met Asn Leu Pro Ser Thr Leu Phe Pro Ser Ile Asp Leu  
 35 40 45  
 Ala Tyr Ile Glu Asp Lys Tyr Ser Asp Ile Ser Arg Gln Arg Leu Phe  
 50 55 60  
 Gly Ser Trp Ser Ser Thr Lys Gly Leu Lys Leu Lys Asn Asp Ile Pro  
 65 70 75  
 Asp Pro Pro Tyr Asn Tyr Asn Asp Thr Lys Val Asp Asp Arg Thr Phe  
 80 85 90  
 Glu Leu Leu Gln Ile Trp Val Ser Thr Asn Pro Arg Gly Ala Glu Arg  
 100 105 110  
 Leu Pro Pro Asp Ile Val Thr Pro Glu Ser Asp Phe Tyr Leu Arg Arg

115 120 125

Leu Trp Gly Asp Pro Asn Glu Asp Leu Thr Val Lys Gln Arg Tyr Leu  
 130 135 140  
 Val Thr Phe Thr Val Gly Tyr Asp Gln Arg Lys Asn Ile Asp Thr Val  
 145 150 155  
 Leu Lys Lys Phe Ser Asp Asn Phe Ser Ile Met Leu Phe His Tyr Asp  
 160 165 170  
 Gly Arg Ala Ser Glu Trp Glu Phe Glu Trp Ser Lys Arg Ala Ile  
 175 180 185  
 His Val Ser Ile Arg Lys Gln Thr Lys Trp Trp Tyr Ala Lys Arg Phe  
 190 195 200  
 Leu His Pro Asp Ile Val Ala Pro Tyr Glu Tyr Ile Phe Ile Trp Asp  
 205 210 215  
 Glu Asp Leu Gly Val Glu His Phe Asp Ser Glu Lys Tyr Leu Ala Val  
 220 225 230  
 Val Lys Lys His Gly Leu Glu Ile Ser Gln Pro Gly Leu Glu Pro Tyr  
 235 240  
 Glu Gly Leu Thr Trp Glu Met Thr Lys Lys Arg Asp Thr Glu Val  
 245 250 255  
 His Lys His Ala Glu Glu Arg Asn Gly Trp Cys Thr Asp Pro Asn Leu  
 260 265 270  
 Pro Pro Cys Ala Ala Phe Val Glu Ile Met Ala Pro Val Phe Ser Arg  
 275 280 285  
 Lys Ala Trp Arg Cys Val Trp His Met Ile Gln Asn Asp Leu Ile His  
 290 295 300  
 Gly Trp Gly Leu Asp Phe Ala Val Arg Lys Cys Val Gln Asn Ala His  
 305 310 315  
 Glu Lys Ile Gly Val Val Asp Ala Gln Trp Ile Ile His Gln Gly Val  
 320 325 330  
 Pro Ser Leu Gly Asn Gln Gly Gln Pro Glu Gln Gly Lys Gln Pro Trp  
 335 340 345  
 Glu Gly Val Arg Glu Arg Cys Arg Arg Glu Trp Thr Met Phe Gln Asp  
 350 355 360  
 Arg Leu Asp Asp Ala Glu Lys Ala Tyr Phe Glu Ala Ser Ala His Lys  
 365 370 375  
 Asn Ala Ser Ser Arg Pro His Gly Asn Trp Val  
 380 385 390  
 405 410 415

&lt;210&gt; 453

&lt;211&gt; 405

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..405:Ceres Seq. ID 2048116

&lt;400&gt;453

Met Met Arg Leu Leu Leu Thr Ser Phe Phe Gly Val Ile Val Gly Phe  
 1 5 10 15  
 Leu Met Gly Ile Thr Pro Thr Thr Leu Thr Leu Thr Lys Met Asn Leu  
 20 25 30  
 Pro Ser Thr Leu Phe Pro Ser Ile Asp Leu Ala Tyr Ile Glu Asp Lys  
 35 40 45  
 Tyr Ser Asp Ile Ser Arg Gln Arg Leu Phe Gly Ser Trp Ser Ser Thr  
 50 55 60  
 Lys Gly Leu Lys Leu Lys Asn Asp Ile Pro Asp Pro Pro Tyr Asn Tyr  
 65 70 75  
 Asn Asp Thr Lys Val Asp Asp Arg Thr Phe Glu Leu Gln Ile Trp  
 80 85 90 95

Val Ser Thr Asn Pro Arg Gly Ala Glu Arg Leu Pro Pro Asp Ile Val  
100 105 110  
Thr Pro Glu Ser Asp Phe Tyr Leu Arg Arg Leu Trp Gly Asp Pro Asn  
115 120 125  
Glu Asp Leu Thr Val Lys Lys Gln Arg Tyr Leu Val Thr Phe Thr Val Gly  
130 135 140  
Tyr Asp Gln Arg Lys Asn Ile Asp Thr Val Leu Lys Lys Phe Ser Asp  
145 150 155  
Asn Phe Ser Ile Met Leu Phe His Tyr Asp Gly Arg Ala Ser Glu Trp  
165 170 175  
Glu Glu Phe Glu Trp Ser Lys Arg Ala Ile His Val Ser Ile Arg Lys  
180 185 190  
Gln Thr Lys Trp Tyr Ala Lys Arg Phe Leu His Pro Asp Ile Val  
195 200 205  
Ala Pro Tyr Glu Tyr Ile Phe Ile Trp Asp Glu Asp Leu Gly Val Glu  
210 215 220  
His Phe Asp Ser Glu Lys Tyr Leu Ala Val Lys Lys His Gly Leu  
225 230 235  
Glu Ile Ser Gln Pro Gly Leu Glu Pro Tyr Glu Gly Leu Thr Trp Glu  
245 250 255  
Met Thr Lys Lys Arg Asp Thr Glu Val His Lys His Ala Glu Glu  
260 265 270  
Arg Asn Gly Trp Cys Thr Asp Pro Asn Leu Pro Pro Cys Ala Ala Phe  
275 280 285  
Val Glu Ile Met Ala Pro Val Phe Ser Arg Lys Ala Trp Arg Cys Val  
290 295 300  
Trp His Met Ile Gln Asn Asp Leu Ile His Gly Trp Gly Leu Asp Phe  
305 310 315  
Ala Val Arg Lys Cys Val Gln Asn Ala His Glu Lys Ile Gly Val Val  
325 330 335  
Asp Ala Gln Trp Ile Ile His Gln Gly Val Pro Ser Leu Gly Asn Gln  
340 345 350  
Gly Gln Pro Glu Gln Gly Lys Gln Pro Trp Glu Gly Val Arg Glu Arg  
355 360 365  
Cys Arg Arg Glu Trp Thr Met Phe Gln Asp Arg Leu Asp Asp Ala Glu  
370 375 380  
Lys Ala Tyr Phe Glu Ala Ser Ala His Lys Asn Ala Ser Ser Arg Pro  
385 390 395  
His Gly Asn Trp Val 405

&lt;210&gt; 454

&lt;211&gt; 404

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..404:Ceres Seq. ID 2048117

&lt;400&gt;454

Met Arg Leu Leu Leu Thr Ser Phe Phe Gly Val Ile Val Gly Phe Leu  
5 10 15  
Met Gly Ile Thr Phe Pro Thr Leu Thr Leu Thr Lys Met Asn Leu Pro  
20 25 30  
Ser Thr Leu Phe Pro Ser Ile Asp Leu Ala Tyr Ile Glu Asp Lys Tyr  
35 40 45  
Ser Asp Ile Ser Arg Gln Arg Leu Phe Gly Ser Trp Ser Ser Thr Lys  
50 55 60  
Gly Leu Lys Leu Lys Asn Asp Ile Pro Asp Pro Pro Tyr Asn Tyr Asn

65 70 75 80  
Asp Thr Lys Val Asp Asp Arg Thr Phe Glu Leu Leu Gln Ile Trp Val  
85 90 95  
Ser Thr Asn Pro Arg Gly Ala Glu Arg Leu Pro Pro Asp Ile Val Thr  
100 105 110  
Pro Glu Ser Asp Phe Tyr Leu Arg Arg Leu Trp Gly Asp Pro Asn Glu  
115 120 125  
Asp Leu Thr Val Lys Gln Arg Tyr Leu Val Thr Phe Thr Val Gly Tyr  
130 135 140  
Asp Gln Arg Lys Asn Ile Asp Thr Val Leu Lys Lys Phe Ser Asp Asn  
145 150 155  
Phe Ser Ile Met Leu Phe His Tyr Asp Gly Arg Ala Ser Glu Trp Glu  
165 170 175  
Glu Phe Glu Trp Ser Lys Arg Ala Ile His Val Ser Ile Arg Lys Gln  
180 185 190  
Thr Lys Trp Trp Tyr Ala Lys Arg Phe Leu His Pro Asp Ile Val Ala  
195 200 205  
Pro Tyr Glu Tyr Ile Phe Ile Trp Asp Glu Asp Leu Gly Val Glu His  
210 215 220  
Phe Asp Ser Glu Lys Tyr Leu Ala Val Lys Lys His Gly Leu Glu  
225 230 235  
Ile Ser Gln Pro Gly Leu Glu Pro Tyr Glu Gly Leu Thr Trp Glu Met  
245 250 255  
Thr Lys Lys Arg Asp Thr Glu Val His Lys His Ala Glu Glu Arg  
260 265 270  
Asn Gly Trp Cys Thr Asp Pro Asn Leu Pro Pro Cys Ala Ala Phe Val  
275 280 285  
Glu Ile Met Ala Pro Val Phe Ser Arg Lys Ala Trp Arg Cys Val Trp  
290 300 305  
His Met Ile Gln Asn Asp Leu Ile His Gly Trp Gly Leu Asp Phe Ala  
310 315 320  
Val Arg Lys Cys Val Gln Asn Ala His Glu Lys Ile Gly Val Val Asp  
325 330 335  
Ala Gln Trp Ile Ile His Gln Gly Val Pro Ser Leu Gly Asn Gln  
340 345 350  
Gln Pro Glu Gln Gly Lys Gln Pro Trp Glu Gly Val Arg Glu Arg Cys  
355 360 365  
Arg Arg Glu Trp Thr Met Phe Gln Asp Arg Leu Asp Asp Ala Glu Lys  
370 375 380  
Ala Tyr Phe Glu Ala Ser Ala His Lys Asn Ala Ser Ser Arg Pro His  
385 390 395  
Gly Asn Trp Val 400

&lt;210&gt; 455

&lt;211&gt; 649

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..649:Ceres Seq. ID 2048271

&lt;400&gt;455

ggaattctc actctctcta tctctcttag ccagctctga tcattctgat tttttccgg  
tgaagaagga ggagaacat ggtgtacgt atcagactgt cagatttgg atcaaaaat  
cgccatttt ttgggttat ggtctgtgat agcagatctc caagacggy gaagatctt  
ggagtttag ttacttcaa tctcttgca gccagagacy gtgtaagag gatggcttc  
aagttgabc gaattaagta ctggttatct gttgtgctc agccatcaga cccggtcaa  
cgtctctact tcagatccgg ttacttctc cctctccaa tgggtgctat gggacgtaaa

ggtagagcac gagacacacg cccagttgat ccgatgactg gtcgtatgt ggcagcagag 420  
aataaacag ttaagccaa tgataacag cctaaggag aggtacaca agccaagagt 480  
gcagtattca ttgacttct gcaacgtag cttttcaagt tcaatttgt gtgagtata 540  
ttgtaatg cagattaga caactgact gtcttccttg ttggcgata aacggcaagg 600  
tgtttgcac ttttgcaga aacggcacat atttgcatt gggatattt 649

<210> 456

<211> 135

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..135:Ceres Seq. ID 2048272

<400>456

Met Val Val Arg Ile Arg Leu Ser Arg Phe Gly Cys Lys Asn Arg Pro 1  
1 5 10 15  
Phe Phe Arg Val Met Ala Ala Asp Ser Arg Ser Pro Arg Asp Gly Lys 20  
20 25 30  
His Leu Glu Val Leu Gly Tyr Phe Asn Pro Leu Pro Gly Gln Asp Gly 35  
35 40 45  
Gly Lys Arg Met Gly Leu Lys Phe Asp Arg Ile Lys Tyr Trp Leu Ser 50  
50 55 60  
Val Gly Ala Gln Pro Ser Asp Pro Val Gln Arg Leu Leu Phe Arg Ser 65  
65 70 75  
Gly Leu Leu Pro Pro Pro Met Val Ala Met Gly Arg Lys Gly Gly 85  
85 90 95  
Ala Arg Asp Thr Arg Pro Val Asp Pro Met Thr Gly Arg Tyr Val Asp 100  
100 105 110  
Ala Glu Asn Lys Thr Val Asn Ala Asn Asp Asn Gln Pro Lys Glu Glu 115  
115 120 125  
Asp Thr Glu Ala Lys Ser Ala 130  
130

<210> 457

<211> 115

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..115:Ceres Seq. ID 2048273

<400>457

Met Ala Ala Asp Ser Arg Ser Pro Arg Asp Gly Lys His Leu Glu Val 1  
1 5 10 15  
Leu Gly Tyr Phe Asn Pro Leu Pro Gly Gln Asp Gly Gly Lys Arg Met 20  
20 25 30  
Gly Leu Lys Phe Asp Arg Ile Lys Tyr Trp Leu Ser Val Gly Ala Gln 35  
35 40 45  
Pro Ser Asp Pro Val Gln Arg Leu Leu Phe Arg Ser Gly Leu Leu Pro 50  
50 55 60  
Pro Pro Pro Met Val Ala Met Gly Arg Lys Gly Gly Ala Arg Asp Thr 65  
65 70 75  
Arg Pro Val Asp Pro Met Thr Gly Arg Tyr Val Asp Ala Glu Asn Lys 85  
85 90 95  
Thr Val Asn Ala Asn Asp Asn Gln Pro Lys Glu Glu Asp Thr Glu Ala 100  
100 105 110

Lys Ser Ala 115

<210> 458

<211> 84

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..84:Ceres Seq. ID 2048274

<400>458

Met Gly Leu Lys Phe Asp Arg Ile Lys Tyr Trp Leu Ser Val Gly Ala 1  
1 5 10 15  
Gln Pro Ser Asp Pro Val Gln Arg Leu Leu Phe Arg Ser Gly Leu Leu 20  
20 25 30  
Pro Pro Pro Pro Met Val Ala Met Gly Arg Lys Gly Gly Ala Arg Asp 35  
35 40 45  
Thr Arg Pro Val Asp Pro Met Thr Gly Arg Tyr Val Asp Ala Glu Asn 50  
50 55 60  
Lys Thr Val Asn Ala Asn Asp Asn Gln Pro Lys Glu Glu Asp Thr Glu 65  
65 70 75  
Ala Lys Ser Ala 80

<210> 459

<211> 285

<212> DNA

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..285:Ceres Seq. ID 2048331

<400>459

atgaactgt gtaatcacca tactgggaag ttgtttcgg agagtcgca agatcagact 60  
gagcaggtac tcaaaacat gggggagata ttgaaageta gtgtgtcga ttattctcg 120  
gtgtggaaga caacaatcat gtgtgtgat ttgtgtgat tcaagcagt gaacagata 180  
tatgccaaat acttccagc tcttttcca gcagatoga cgtatcaagt tgcagcttg 240  
ccttcaacg ccaagatcga gattgaatgt attgcaacac tctag 285

<210> 460

<211> 94

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..94:Ceres Seq. ID 2048332

<400>460

Met Lys Leu Cys Asn His His Thr Gly Lys Phe Val Ser Glu Ser Val 1  
1 5 10 15  
Glu Asp Gln Thr Gln Gln Val Leu Lys Asn Met Gly Glu Ile Leu Lys 20  
20 25 30  
Ala Ser Gly Ala Asp Tyr Ser Ser Val Val Lys Thr Thr Ile Met Leu 35  
35 40 45

Ala Asp Leu Ala Asp Phe Lys Thr Val Asn Glu Ile Tyr Ala Lys Tyr  
50 55 60  
Phe Pro Ala Pro Ser Pro Ala Arg Ser Thr Tyr Gln Val Ala Ala Leu  
65 70 75 80  
Pro Leu Asn Ala Lys Ile Glu Ile Gln Cys Ile Ala Thr Leu  
85 90

&lt;210&gt; 461

&lt;211&gt; 68

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..66:Ceres Seq. ID 2048333

&lt;400&gt;461

Met Gly Glu Ile Leu Lys Ala Ser Gly Ala Asp Tyr Ser Ser Val Val  
1 5 10 15  
Lys Thr Thr Ile Met Leu Ala Asp Leu Ala Asp Phe Lys Thr Val Asn  
20 25 30  
Glu Ile Tyr Ala Lys Tyr Phe Pro Ala Pro Ser Pro Ala Arg Ser Thr  
35 40 45  
Tyr Gln Val Ala Ala Leu Pro Leu Asn Ala Lys Ile Glu Ile Glu Cys  
50 55 60  
Ile Ala Thr Leu  
65

&lt;210&gt; 462

&lt;211&gt; 48

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..48:Ceres Seq. ID 2048334

&lt;400&gt;462

Met Leu Ala Asp Leu Ala Asp Phe Lys Thr Val Asn Glu Ile Tyr Ala  
1 5 10 15  
Lys Tyr Phe Pro Ala Pro Ser Pro Ala Arg Ser Thr Tyr Gln Val Ala  
20 25 30  
Ala Leu Pro Leu Asn Ala Lys Ile Glu Ile Glu Cys Ile Ala Thr Leu  
35 40 45

&lt;210&gt; 463

&lt;211&gt; 627

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..627:Ceres Seq. ID 2048466

&lt;400&gt;463

atgatggaca agaataaaaa ccctagtitt ttgtctctc gaagttcgt ggcactctct  
60  
gctcgagcaa gatctctac cgatacaatg attatctcag aggtacaocg caaagaatc  
120  
tgcgaagacc tcttcaaga aggaatttgc tttycgaaga aggaatttcaa ttctgtcaag  
180

catcgttgta tcatgtccc caactaca gtattaagc ttatgcagag ttcaaatcc  
aagagtatg ttaggagac attgctcgg atgcattact attggtttt gaccaatga  
300  
ggatcgagt ttctgagaac ttacttaat ttccatctg atgtgttcc tgcactttg  
360  
aagagtcag ctaagcttgg tggctcca ttgttggcc caactgttga tgcctcaga  
420  
ggactctcc atgaagaggg agacgtccc aggtttggg accgtgatgg gtacgtgca  
480  
ggctctcag ctggtgttga gtttgaggt gaaaagggtg gagctccgc agattaccag  
540  
ccattctcc aaggaagtgg ccgtgtttt ggcgtgtgtg ctggtgcta cagcgcagt  
600  
gcaccatcg gttcagggtt gcttga  
627

&lt;210&gt; 464

&lt;211&gt; 208

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..208:Ceres Seq. ID 2048467

&lt;400&gt;464

Met Met Asp Lys Asn Glu Asn Pro Ser Phe Leu Leu Ser Arg Ser Ser  
1 5 10 15  
Val Arg Leu Ser Ala Arg Ala Arg Ser Pro Thr Asp Thr Met Ile Ile  
20 25 30  
Ser Glu Ala Asn Arg Lys Glu Ile Cys Lys Tyr Leu Phe Lys Glu Gly  
35 40 45  
Val Cys Phe Ala Lys Lys Asp Phe Asn Leu Ala Lys His Pro Leu Ile  
50 55 60  
Asp Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln Ser Phe Lys Ser  
65 70 75 80  
Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His Tyr Tyr Trp Phe  
85 90 95  
Leu Thr Asn Glu Gly Ile Glu Phe Leu Arg Thr Tyr Leu Asn Leu Pro  
100 105 110  
Ser Asp Val Val Pro Ala Thr Leu Lys Lys Ser Ala Lys Pro Gly Gly  
115 120 125  
Arg Pro Phe Gly Gly Pro Pro Gly Asp Arg Ser Arg Gly Pro Arg His  
130 135 140  
Glu Gly Gly Asp Arg Pro Arg Phe Gly Asp Arg Asp Gly Tyr Arg Ala  
145 150 155 160  
Gly Pro Arg Ala Gly Gly Glu Phe Gly Gly Lys Gly Gly Ala Pro  
165 170 175  
Ala Asp Tyr Gln Pro Ser Phe Gln Gly Ser Gly Arg Gly Phe Gly Arg  
180 185 190  
Gly Ala Gly Gly Tyr Ser Ala Ala Pro Ser Gly Ser Gly Leu Pro  
195 200 205

&lt;210&gt; 465

&lt;211&gt; 207

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..207:Ceres Seq. ID 2048468

&lt;400&gt;465

Met Asp Lys Asn Glu Asn Pro Ser Phe Leu Leu Ser Arg Ser Ser Val  
1 5 10 15  
Arg Leu Ser Ala Arg Ala Arg Ser Pro Thr Asp Thr Met Ile Ile Ser

Glu Ala Asn Arg Lys Glu Ile Cys Lys Tyr Leu Phe Lys Glu Gly Val 30  
 35  
 Cys Phe Ala Lys Lys Asp Phe Asn Leu Ala Lys His Pro Leu Ile Asp 45  
 50  
 Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln Ser Phe Lys Ser Lys 60  
 65  
 Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His Tyr Tyr Trp Phe Leu 80  
 85  
 Thr Asn Glu Gly Ile Glu Phe Leu Arg Thr Tyr Leu Asn Leu Pro Ser 100  
 110  
 Asp Val Val Pro Ala Thr Leu Lys Lys Ser Ala Lys Pro Gly Gly Arg 115  
 120  
 Pro Phe Gly Gly Pro Pro Gly Asp Arg Ser Arg Gly Pro Arg His Glu 130  
 135  
 Gly Gly Asp Arg Pro Arg Phe Gly Asp Arg Asp Gly Tyr Arg Ala Gly 140  
 145  
 Pro Arg Ala Gly Gly Glu Phe Gly Gly Glu Lys Gly Ala Pro Ala 150  
 155  
 Asp Tyr Gln Pro Ser Phe Gln Gly Ser Gly Arg Gly Phe Gly Arg Gly 165  
 170  
 Ala Gly Gly Tyr Ser Ala Ala Ala Pro Ser Gly Ser Gly Leu Pro 180  
 185  
 190  
 195

&lt;210&gt; 466

&lt;211&gt; 179

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..179:Ceres Seq. ID 2048469

Met Ile Ile Ser Glu Ala Asn Arg Lys Glu Ile Cys Lys Tyr Leu Phe 1  
 5  
 Lys Glu Gly Val Cys Phe Ala Lys Lys Asp Phe Asn Leu Ala Lys His 10  
 15  
 Pro Leu Ile Asp Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln Ser 20  
 25  
 Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His Tyr 30  
 35  
 Tyr Trp Phe Leu Thr Asn Glu Gly Ile Glu Phe Leu Arg Thr Tyr Leu 40  
 45  
 50  
 55  
 60  
 65  
 Asn Leu Pro Ser Asp Val Val Pro Ala Thr Leu Lys Ser Ala Lys 70  
 75  
 Pro Gly Gly Arg Pro Phe Gly Gly Pro Pro Gly Asp Arg Ser Arg Gly 80  
 85  
 90  
 95  
 100  
 105  
 110  
 115  
 120  
 125  
 Tyr Arg Ala Gly Pro Arg Ala Gly Gly Glu Phe Gly Glu Lys Gly 130  
 135  
 Gly Ala Pro Ala Asp Tyr Gln Pro Ser Phe Gln Gly Ser Gly Arg Gly 140  
 145  
 Phe Gly Arg Gly Ala Gly Tyr Ser Ala Ala Ala Pro Ser Gly Ser 150  
 155  
 160  
 165  
 170  
 175  
 Gly Leu Pro

&lt;210&gt; 467

<211> 585  
 <212> DNA  
 <213> Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..585:Ceres Seq. ID 2050485

<400>467  
 attctcgaa agtctcatt ctgcacccc aatctggga ttagggttaa aagaaccatt 60  
 ttattctcg tgcgcaaca acaattccg atcgaaaag gaagaagaga tgaatggc 120  
 gtgagaagg gttacagtg aaatcagtg gaagaagtg acgagcttc caggctatat 180  
 caacgact ttttcattg agaccgtga gacctctg agagaggac tggataacta 240  
 caacgaata tactataga ccagctcgt tgatctatc ctctcatct gctttacg 300  
 catgcttc tcttaoctg tgcctccc taatgagct cgcactctg agcatcaga 360  
 gaactaag gagacagtg gtcattgac tctgggaat cgttcgact tgcagatgat 420  
 tttaggggg ttgctgga atctttct gcttgatgt gacgacgt caagattgt 480  
 gttctatgt ttgctttct tgaatttcc tggataatg tgnctaaag gaaaacctt 540  
 ctttgaatt acattccatg atagtcaata attgaagcat catga 585

&lt;210&gt; 468

&lt;211&gt; 128

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..128:Ceres Seq. ID 2050486

<400>468  
 Ser Ser Glu Ser Leu Ile Ser Arg Ser Pro Ile Arg Gly Leu Gly Leu 1  
 5  
 Lys Glu Pro Phe Leu Phe Ser Ser Arg Asn Asn Lys Ser Arg Ser Lys 10  
 15  
 Lys Glu Glu Glu Ile Glu Met Ala Leu Arg Arg Val Tyr Ser Glu Ile 20  
 25  
 30  
 35  
 40  
 45  
 Arg Gly Lys Lys Val Thr Glu Leu Pro Gly Tyr Ile Lys Ser Thr Phe 50  
 55  
 Ser Met Glu Thr Val Lys Thr Ser Val Lys Arg Gly Leu Asp Asn Tyr 60  
 65  
 70  
 75  
 80  
 85  
 90  
 95  
 100  
 105  
 110  
 115  
 120  
 125

&lt;210&gt; 469

&lt;211&gt; 90

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..90:Ceres Seq. ID 2050487

&lt;400&gt;469

Met Ala Leu Arg Arg Val Tyr Ser Glu Ile Arg Gly Lys Lys Val Thr

1  
Glu Leu Pro Gly Tyr Ile Lys Ser Thr Phe Ser Met Glu Thr Val Lys 15  
20 25 30  
Thr Ser Val Lys Arg Gly Leu Asp Asn Tyr Asn Glu Lys Tyr Ile Gln 35  
40 45  
Thr Ser Ser Val Asp Pro Ile Leu His Ile Cys Phe Tyr Gly Met Ala 50  
55 60  
Phe Ser Tyr Leu Val Ala Leu Pro Asn Glu Arg Arg His Leu Glu His 65  
70 75 80  
Gln Gln His Ala Lys Glu His Gly Gly His 85  
90

&lt;210&gt; 470

&lt;211&gt; 516

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..516:Ceres Seq. ID 2050708

&lt;400&gt;470

caaccaaac aaacataaa aaacaagtgg aagctttaaa acgagaggga gagagcaaaa 60  
atggcgagct cggagacgta cgtgacgga gttccgctaa agggatggc cgagaaatac 120  
tacaagggt ggaagaaaga gaacatgac ttcctgatg ctatcgcca caacatccaa 180  
aatgtaccg ttacagaaag cgacatgac tctcagggg ctatcagag ttggaactac 240  
acatggatg gaaaggagg ggtgtcaag gagagagag agatagaca tgaacccaa 300  
acgttgagc taagagagct tggaggtcac gtgagggag agtccaagt gtagcagtc 360  
gtctaccaat tcatcccaa atctggagt acctgcacg gcaaatcac ttaatatgg 420  
gagaagcca acgatgatic ccagaaaca agcgggtaca tgaatttgt caagagcttg 480  
gtgtctgaca tgggaacca cgttagcaaa acttaa 516

&lt;210&gt; 471

&lt;211&gt; 151

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..151:Ceres Seq. ID 2050709

&lt;400&gt;471

Met Ala Thr Ser Gly Thr Tyr Val Thr Glu Val Pro Leu Lys Gly Ser 1  
5 10 15  
Ala Glu Lys Tyr Tyr Lys Arg Trp Lys Asn Glu Asn His Val Phe Pro 20  
25 30  
Asp Ala Ile Gly His His Ile Gln Asn Val Thr Val His Glu Gly Glu 35  
40 45  
His Asp Ser His Gly Ser Ile Arg Ser Trp Asn Tyr Trp Asp Gly 50  
55 60  
Lys Glu Glu Val Phe Lys Glu Arg Arg Glu Ile Asp Asp Glu Thr Lys 65  
70 75 80  
Thr Leu Thr Leu Arg Gly Leu Glu Gly His Val Met Glu Gln Leu Lys 85  
90 95  
Val Tyr Asp Val Val Tyr Gln Phe Ile Pro Lys Ser Glu Asp Thr Cys 100  
105 110  
Ile Gly Lys Ile Thr Leu Ile Trp Glu Lys Arg Asn Asp Asp Ser Pro 115  
120 125  
Glu Pro Ser Gly Tyr Met Lys Phe Val Lys Ser Leu Val Ala Asp Met

130 135 140  
Gly Asn His Val Ser Lys Thr 145  
150

&lt;210&gt; 472

&lt;211&gt; 1730

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..1730:Ceres Seq. ID 2050901

&lt;400&gt;472

atcggtgtt actctttt aaaggaacca ttgtatatca tggaggctt gacttacttg 60  
aagctattt gcttccaaag gaattcaagg ttcttttca cgtgaattct ctggaatag 120  
ctatggagt acttcagaac acctgatac cctacgaaa cgcacaatt gctctccag 180  
acctgtccc tgaagtaaa aaacgaacc aaacagagag cattgttca tataaagct 240  
cgagaatcac cgaataagc ctctctcta gcaggttttg gaagatcata taccgtaca 300  
ggcagttgct tetaaacaac atctagaat ctctgtagt gggcttgic ttaggacta 360  
ttacacttaa tatcggaact ggcaagaag gaacaggaa acgatttggc ctttcgat 420  
tcacctaac attctcttc tctccacta cccaaacct tccatattc attgagaac 480  
gacctattt ttcccgaaa acctcaagcg gacttcacg actctctct cacattctg 540  
caaacactt ggtttcttg ccaacttgc tactatcgc aatcattct tctgtctac 600  
tctattctt tgttagact tcttttcat ggaagctct cgtctactt gtctcgtaa 660  
ttgtgatct tgtcctaag gtaactctt ttgtacttt ttgagctct ctgcacata 720  
actacatgc tggacactc taagtacca ttctcttgc ggtttcttc ttgtcttg 780  
gtacttact cttaaaag agtcttcca agtactgct ctcatgtac ttcttctaa 840  
ttacaagta tgggttgac gacttttga taatagta ctcgtctcg caacaagt 900  
ctcgtctcg gtttagaaa gcttcttga atagctctt agttactga ggtgagctg 960  
tagacaaga tgggttcat gagagaca gatatatgt gtgggagc cgtctattgg 1020  
tctccctt caacttag aaacaagaa tctgttctt gagaagcaa aaacctat 1080  
tctccctt tgcacaagt cagtgcctc ttgcagcaa agtcaacca caagacgt 1140  
ggaattacc gagtataat tgtatcttg ttctctctt tatcagatta tctcagaaa 1200  
caatcgaaa gagatttca aatactttt caaagtagt ggaacagag agagtatgt 1260  
ttgtcaaaa ggaattcaa ctccaaagc atccgttga tgaatgaca aactgcaag 1320  
tgattaagt catcagagt ttcaagtcca aggttagct tagggagaca ttgtctgga 1380  
tgcattata ttggttctg actaatga gaaatgagt ttgagaact tatctaac 1440  
ttcttccga tgtttctct gtaatttga aagaatgac taagccggt ggtgtccct 1500  
ttgtgtccc acctgtgat cgccaaagag gaccacctg ctctgtaga gaccttcca 1560  
gatttgta cagtgatga taccgtgag gccacgttg tggtagtag aaggttgag 1620  
ctccagcta ttccagcgt tcttccaa gagggttg taggcctgt ttggcctg 1680  
gtgcagcgt ttaccgtga gcagcaact ctggttcagg gtcccttga 1730

&lt;210&gt; 473

&lt;211&gt; 575

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..575:Ceres Seq. ID 2050902

&lt;400&gt;473

Arg Val Leu Leu Leu Ser Lys Gly Thr Ile Val Tyr His Gly Arg Leu 1  
5 10 15  
Asp Leu Leu Glu Ala Phe Leu Leu Ser Lys Gly Phe Thr Val Pro Ser 20  
25 30  
Gln Leu Asn Ser Leu Glu Tyr Ala Met Glu Ile Leu Gln Asn Ile Arg

35 Asp Pro Tyr Glu Asn Ala Asn Ile Ala Leu Pro Asp His Cys Pro Glu 45  
 50 Ser Lys Lys Glu Asn Gln Lys Ser Ile Val Arg Tyr Lys Ser Ser 60  
 65 Arg Ile Thr Glu Ile Ser Leu Ser Ser Arg Phe Trp Lys Ile Ile 75  
 85 Tyr Arg Thr Arg Glu Leu Leu Thr Asn Ile Leu Glu Ser Leu Val 100  
 105 Val Gly Leu Val Leu Gly Thr Ile Tyr Leu Asn Ile Gly Thr Gly Lys 110  
 115 Glu Gly Ile Arg Lys Arg Phe Gly Leu Phe Ala Phe Thr Leu Thr Phe 125  
 130 Leu Leu Ser Ser Thr Thr Gln Thr Leu Pro Ile Phe Ile Asp Glu Arg 140  
 145 Pro Ile Leu Leu Arg Glu Thr Ser Ser Gly Leu Tyr Arg Leu Ser Ser 155  
 165 His Ile Leu Ala Asn Thr Leu Val Phe Leu Pro Tyr Leu Leu Ile 170  
 180 Ala Ile Ile Tyr Ser Val Ser Leu Tyr Phe Leu Val Gly Leu Cys Phe 185  
 195 Ser Trp Gln Ala Leu Ala Tyr Phe Val Leu Val Ile Trp Ile Ile Val 200  
 210 Leu Met Ala Asn Ser Phe Val Leu Phe Leu Ser Ser Leu Ala Pro Asn 215  
 225 Tyr Ile Ala Gly Thr Ser Ser Val Thr Ile Leu Leu Ala Ala Phe Phe 230  
 245 Leu Phe Ser Gly Tyr Phe Ile Ser Lys Glu Ser Leu Pro Lys Tyr Trp 250  
 260 Leu Phe Met Tyr Phe Phe Ser Met Tyr Lys Tyr Ala Leu Asp Ala Leu 265  
 275 Leu Ile Asn Glu Tyr Ser Cys Leu His Asn Lys Cys Leu Val Trp Phe 280  
 290 Glu Glu Ala Ser Val Asn Ser Cys Leu Val Thr Gly Asp Val Leu 295  
 305 Asp Lys Asn Gly Leu His Glu Arg Gln Arg Ile Tyr Val Leu Gly Arg 310  
 325 Pro Ser Leu Val Ser Gly Phe Lys Leu Glu Lys Gln Gly Ile Arg Phe 330  
 345 Leu Arg Ser Lys Lys Thr Leu Leu Pro Phe Ala Gln Ala Gln Cys 350  
 360 Leu Phe Ala Ala Lys Ala Gln Pro Thr Arg Arg Trp Asn Tyr Pro Ser 365  
 370 Ile Asn Cys Asp Leu Val Arg Pro Phe Ile Arg Leu Ser Gln Arg Thr 375  
 385 Ile Ala Glu Arg Ser Ala Asn Thr Phe Ser Lys Leu Val Glu Pro Glu 390  
 405 Gly Val Cys Phe Ala Lys Lys Asp Phe Asn Leu Pro Lys His Pro Leu 410  
 420 Ile Asp Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln Ser Phe Lys 425  
 435 Ser Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His Tyr Tyr Trp 440  
 450 Phe Leu Thr Asn Glu Gly Ile Glu Phe Leu Arg Thr Tyr Leu Asn Leu 445  
 465 Pro Ser Asp Val Val Pro Ala Thr Leu Lys Lys Ser Ala Lys Pro Gly 470  
 485 Gly Arg Pro Phe Gly Gly Pro Pro Gly Asp Arg Gln Arg Gly Pro Pro 485  
 500 Arg Ser Asp Gly Asp Arg Pro Arg Phe Gly Asp Arg Asp Gly Tyr Arg 510  
 515

Gly Gly Pro Arg Gly Gly Asp Glu Lys Gly Gly Ala Pro Ala Asp Phe 530  
 535 Gln Pro Ser Phe Gln Gly Gly Gly Arg Pro Gly Phe Gly Arg Gly 540  
 545 Ala Gly Gly Tyr Ser Ala Ala Ala Pro Ser Gly Ser Gly Phe Pro 550  
 555 <210> 474  
 <211> 535  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <220>  
 <223> any n or Xaa = unknown  
 <223> LOCATION: 1..535:Ceres Seq. ID 2050903  
 <400> 474  
 1 Met Glu Ile Leu Gln Asn Ile Arg Asp Pro Tyr Glu Asn Ala Asn Ile 5  
 10 Ala Leu Pro Asp His Cys Pro Glu Ser Lys Lys Gln Asn Gln Lys Gln 20  
 25 Ser Ile Val Arg Tyr Lys Ser Ser Arg Ile Thr Glu Ile Ser Leu Leu 30  
 35 Ser Ser Arg Phe Trp Lys Ile Ile Tyr Arg Thr Arg Gln Leu Leu Leu 40  
 45 Thr Asn Ile Leu Glu Ser Leu Val Val Gly Leu Val Leu Gly Thr Ile 50  
 55 Tyr Leu Asn Ile Gly Thr Gly Lys Glu Gly Ile Arg Lys Arg Phe Gly 60  
 65 Leu Phe Ala Phe Thr Leu Thr Phe Leu Leu Ser Ser Thr Thr Gln Thr 85  
 90 Leu Pro Ile Phe Ile Asp Glu Arg Pro Ile Leu Leu Arg Glu Thr Ser 100  
 105 Ser Gly Leu Tyr Arg Leu Ser Ser His Ile Leu Ala Asn Thr Leu Val 110  
 115 Phe Leu Pro Tyr Leu Leu Ile Ala Ile Ile Tyr Ser Val Ser Leu 120  
 125 Tyr Phe Leu Val Gly Leu Cys Phe Ser Trp Gln Ala Leu Ala Tyr Phe 130  
 135 Val Leu Val Ile Trp Ile Ile Val Leu Met Ala Asn Ser Phe Val Leu 140  
 145 Phe Leu Ser Ser Leu Ala Pro Asn Tyr Ile Ala Gly Thr Ser Ser Val 150  
 155 Thr Ile Leu Leu Ala Ala Phe Phe Leu Phe Sor Gly Tyr Phe Ile Ser 160  
 165 Lys Glu Ser Leu Pro Lys Tyr Trp Leu Phe Met Tyr Phe Ser Met 170  
 175 Tyr Lys Tyr Ala Leu Asp Ala Leu Leu Ile Asn Glu Tyr Ser Cys Leu 180  
 185 His Asn Lys Cys Leu Val Trp Phe Glu Ala Ser Val Asn Ser Cys 190  
 195 Leu Val Thr Gly Gly Asp Val Leu Asp Lys Asn Gly Leu His Glu Arg 200  
 205 Gln Arg Ile Tyr Val Leu Gly Arg Pro Ser Leu Val Ser Gly Phe Lys 210  
 215 Leu Glu Lys Gln Gly Ile Arg Phe Leu Arg Ser Lys Lys Thr Leu Leu 220  
 225 Leu Pro Phe Ala Gln Ala Gln Cys Leu Phe Ala Ala Lys Ala Gln Pro 230  
 235 Thr Arg Arg Trp Asn Tyr Pro Ser Ile Asn Cys Asp Leu Val Arg Pro 240  
 245

340 Phe Ile Arg Leu Ser Gln Arg Thr Ile Ala Glu Arg Ser Ala Asn Thr 350  
 355 Phe Ser Lys Leu Val Glu Pro Glu Gly Val Cys Phe Ala Lys Lys Asp 365  
 370 Phe Asn Leu Pro Lys His Pro Leu Ile Asp Val Pro Asn Leu Gln Val 380  
 385 Ile Lys Leu Met Gln Ser Phe Lys Ser Lys Glu Tyr Val Arg Glu Thr 400  
 405 Phe Ala Trp Met His Tyr Tyr Trp Phe Leu Thr Asn Glu Gly Ile Glu 415  
 420 Phe Leu Arg Thr Tyr Leu Asn Leu Ser Ser Asp Val Val Pro Ala Thr 430  
 435 Leu Lys Lys Ser Ala Lys Pro Gly Gly Arg Pro Phe Gly Gly Pro Pro 445  
 450 Gly Asp Arg Gln Arg Gly Pro Pro Arg Ser Asp Gly Asp Arg Pro Arg 460  
 465 Phe Gly Asp Arg Asp Gly Tyr Arg Gly Gly Pro Arg Gly Gly Asp Glu 475  
 485 Lys Gly Gly Ala Pro Ala Asp Phe Gln Pro Ser Phe Gln Gly Gly Gly 490  
 500 Gly Arg Pro Gly Phe Gly Arg Gly Ala Gly Gly Tyr Ser Ala Ala Ala 510  
 515 Pro Ser Gly Ser Gly Phe Pro 520  
 530 535

&lt;210&gt; 475

&lt;211&gt; 350

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..350:Ceres Seq. ID 2050904

&lt;400&gt;475

Met Ala Asn Ser Phe Val Leu Phe Leu Ser Ser Leu Ala Pro Asn Tyr 15  
 1 Ile Ala Gly Thr Ser Ser Val Thr Ile Leu Leu Ala Ala Phe Phe Leu 30  
 20 Phe Ser Gly Tyr Phe Ile Ser Lys Glu Ser Leu Pro Lys Tyr Trp Leu 45  
 35 Phe Met Tyr Phe Phe Ser Met Tyr Lys Tyr Ala Leu Asp Ala Leu Leu 60  
 50 Ile Asn Glu Tyr Ser Cys Leu His Asn Lys Cys Leu Val Trp Phe Glu 75  
 65 Glu Ala Ser Val Asn Ser Cys Leu Val Thr Gly Gly Asp Val Leu Asp 95  
 85 Lys Asn Gly Leu His Glu Arg Gln Arg Ile Tyr Val Leu Gly Arg Pro 110  
 100 Ser Leu Val Ser Gly Phe Lys Leu Glu Lys Gln Gly Ile Arg Phe Leu 125  
 115 Arg Ser Lys Lys Thr Leu Leu Leu Pro Phe Ala Gln Ala Gln Cys Leu 140  
 130 Phe Ala Ala Lys Ala Gln Pro Thr Arg Arg Trp Asn Tyr Pro Ser Ile 155  
 145 Asn Cys Asp Leu Val Arg Pro Phe Ile Arg Leu Ser Gln Arg Thr Ile 175  
 165 Ala Glu Arg Ser Ala Asn Thr Phe Ser Lys Leu Val Glu Pro Glu Gly 185  
 180

Val Cys Phe Ala Lys Lys Asp Phe Asn Leu Pro Lys His Pro Leu Ile 195  
 200 Asp Val Pro Asn Leu Gln Val Ile Lys Leu Met Gln Ser Phe Lys Ser 220  
 210 Lys Glu Tyr Val Arg Glu Thr Phe Ala Trp Met His Tyr Tyr Trp Phe 235  
 225 Leu Thr Asn Glu Gly Ile Glu Phe Leu Arg Thr Tyr Leu Asn Leu Pro 250  
 260 Ser Asp Val Val Pro Ala Thr Leu Lys Lys Ser Ala Lys Pro Gly Gly 270  
 265 Arg Pro Phe Gly Gly Pro Pro Gly Asp Arg Gln Arg Gly Pro Pro Arg 285  
 275 Ser Asp Gly Asp Arg Pro Arg Phe Gly Asp Arg Asp Gly Tyr Arg Gly 300  
 290 Gly Pro Arg Gly Gly Asp Glu Lys Gly Gly Ala Pro Ala Asp Phe Gln 310  
 305 Pro Ser Phe Gln Gly Gly Gly Arg Pro Gly Phe Gly Arg Gly Ala 330  
 325 Gly Gly Tyr Ser Ala Ala Pro Ser Gly Ser Gly Phe Pro 345  
 340

&lt;210&gt; 476

&lt;211&gt; 535

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..535:Ceres Seq. ID 2051325

&lt;400&gt;476

gctacataac tcaagtaaac aaacatacac atataacata tagaagatc gaacacaaa 60  
 gatcgggatc agtatggctc caaagacctc aactacactt gcttattcc ttgtaccacaa 120  
 tatctctc ctcaaccica ttacactgag ttgcgcagac aatactggc caagagacgt 180  
 ttctaaactt tgcacatgct cgaatgttct caactctatc aacttgaac tgggggccc 240  
 agtataggc ccttgatgct ctattctctt tggctaat ttgtctgat ttgtgtttg 300  
 cctttgacc ggcctcaagc tcagctttct tggcatcaac atcgacactc ctattacct 360  
 taacttgctt cttaacgctt gggagggtac cctctcgat ggattcogtt gccacacata 420  
 gctacatat attcatattt tctatctatg tcttgaatt ggtttgctt aacatccaaa 480  
 ttctaaactt aaattcatgt ataagcaat aacagtaaaa attaagactt ttatgt 535

&lt;210&gt; 477

&lt;211&gt; 117

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..117:Ceres Seq. ID 2051326

&lt;400&gt;477

Met Ser Met Ala Pro Lys Thr Ser Thr Thr Leu Ala Leu Phe Leu Val 15  
 1 Thr Asn Ile Leu Phe Leu Asn Leu Ile Thr Leu Ser Cys Ala Asp Asn 30  
 20 Thr Cys Pro Arg Asp Val Leu Lys Leu Ser Thr Cys Ser Asn Val Leu 45  
 35 Asn Leu Ile Asn Leu Lys Leu Gly Ala Pro Ala Met Arg Pro Cys Cys 60  
 55



Ser Ile Leu Phe Gly Leu Ile Asp Leu Asp Val Ala Val Cys Leu Cys  
65 70 75 80  
Thr Ala Leu Lys Leu Ser Leu Leu Gly Ile Thr Ile Asp Thr Pro Ile  
85 90 95  
His Leu Asn Leu Ala Leu Asn Ala Cys Gly Thr Leu Pro Asp Gly  
100 110  
Phe Arg Cys Pro Thr  
115

<210> 478  
<211> 115  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..115:Ceres Seq. ID 2051327

<400>478  
Met Ala Pro Lys Thr Ser Thr Thr Leu Ala Leu Phe Leu Val Thr Asn  
1 5 10 15  
Ile Leu Phe Leu Asn Leu Ile Thr Leu Ser Cys Ala Asp Asn Thr Cys  
20 25 30  
Pro Arg Asp Val Leu Lys Leu Ser Thr Cys Ser Asn Val Leu Asn Leu  
35 40 45  
Ile Asn Leu Lys Leu Gly Ala Pro Ala Met Arg Pro Cys Cys Ser Ile  
50 55 60  
Leu Phe Gly Leu Ile Asp Leu Asp Val Ala Val Cys Leu Cys Thr Ala  
65 70 75 80  
Leu Lys Leu Ser Leu Leu Gly Ile Thr Ile Asp Thr Pro Ile His Leu  
85 90 95  
Asn Leu Ala Leu Asn Ala Cys Gly Thr Leu Pro Asp Gly Phe Arg  
100 105 110  
Cys Pro Thr  
115

<210> 479  
<211> 74  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..74:Ceres Seq. ID 2051328

<400>479  
Met Leu Leu Phe Ala Phe Ala Pro Arg Ser Ser Ala Phe Leu Ala  
1 5 10 15  
Ser Pro Ser Thr Leu Leu Phe Thr Leu Thr Trp Leu Leu Thr Pro Val  
20 25 30  
Glu Val Pro Phe Leu Met Asp Ser Val Ala Gln His Ser Tyr Asn Ile  
35 40 45  
Phe Ile Phe Ser Ile Tyr Val Leu Glu Leu Val Leu Ser Asn Ile Gln  
50 55 60  
Ile Ser Asn Leu Asn Ser Cys Ile Ser Lys  
65 70

<210> 480  
<211> 517

<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..517:Ceres Seq. ID 2051633

<400>480  
cataaccctc ttggaagag tctcaacct tgcagagaaa aagaacaagg aagatcccgg  
aaatggcaa cggcgattgt acgttaagct ctctcccgag cagtgactcg cgaagctccg  
aagacatccg tgcctctaa gcgaacttt tctctccg cggccatga cgaagcttat  
gaagctgcga agtggagaa gataacttat ctgggtatg ctagtgcac tctctagct  
gtctatgtt tatcaagggg ccatcatcac ggcgaagacc ctctgccta tctgcatatg  
cacatcgca acaaggagtt tctctgggt cggatggtc tgttgaggt gaagcacac  
aagagacact gsgctctgcg tggcataat aacgtctct tggttattt gaaggetaa  
aatgtttac cgtatttgtt ctacacgttt gteaagatt tgetactoca atctctttc  
tttgttggg aataaaagt taactacttg ctgggc

<210> 481  
<211> 123  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..123:Ceres Seq. ID 2051634

<400>481  
His Asn Pro Leu Gly Lys Ser Leu Asn Thr Cys Arg Glu Lys Glu Gln  
1 5 10 15  
Gly Arg Ser Arg Lys Met Ala Thr Ala Ile Val Arg Ser Ala Leu Sor  
20 25 30  
Arg Ala Val Thr Arg Ala Ala Pro Lys Thr Ser Val Ala Pro Lys Arg  
35 40 45  
Asn Phe Ser Ser Ala Gly His Asp Asp Ala Tyr Glu Ala Ala Lys  
50 55 60  
Trp Glu Lys Ile Thr Tyr Leu Gly Ile Ala Ser Cys Thr Ala Leu Ala  
65 70 75 80  
Val Tyr Val Leu Ser Lys Gly His His His Gly Glu Asp Pro Ala  
85 90 95  
Tyr Pro His Met His Ile Arg Asn Lys Glu Phe Pro Trp Gly Pro Asp  
100 105 110  
Gly Leu Phe Glu Val Lys His Asn Lys Glu His  
115 120

<210> 482  
<211> 102  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..102:Ceres Seq. ID 2051635

<400>482  
Met Ala Thr Ala Ile Val Arg Ser Ala Leu Ser Arg Ala Val Thr Arg  
1 5 10 15  
Ala Ala Pro Lys Thr Ser Val Ala Pro Lys Arg Asn Phe Ser Ser Ser

20 25 30  
Ala Gly His Asp Ala Tyr Glu Ala Ala Lys Trp Glu Lys Ile Thr  
35 40 45  
Tyr Leu Gly Ile Ala Ser Cys Thr Ala Leu Ala Val Tyr Val Leu Ser  
50 55 60  
Lys Gly His His His Gly Glu Asp Pro Pro Ala Tyr Pro His Met His  
65 70 75 80  
Ile Arg Asn Lys Glu Phe Pro Trp Gly Pro Asp Gly Leu Phe Glu Val  
85 90 95  
Lys His Asn Lys Glu His  
100

<210> 483  
<211> 410  
<212> DNA  
<213> Arabidopsis thaliana  
<220>  
<223> any n or Xaa = unknown  
<223> LOCATION: 1..410:Ceres Seq. ID 2051906

<400>483  
ttctctttt ttacatcta ccaaaattc caaaacacca tgcattttt gctctctga 60  
aatgtctt ttatgtcta attctctct tcaactaaa cgaacaggga aaacaggtt 120  
cggagatcgg tgttactga tggccaaca gcaagaaacc cgaacttaca tactcgacy 180  
atgtgtctcc atgtgtcttt gctggacaga ccaacttatt tctgattagt ttgattcag 240  
aacggacgg accaagaagt agaaacaga agaattgat gtcatttga gttttcttg 300  
tttcttcta gaactctca aggaattctt tgaattgaa gaagactcta aggattgggc 360  
ctttgggccc ttgtacata ttgtgaaat aacatgtaa cttttttac 410

<210> 484  
<211> 75  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..75:Ceres Seq. ID 2051907

<400>484  
Ser Phe Phe Phe Thr Ser Thr Gln Asn Ser Lys Thr Pro Ser Ile Phe  
1 5 10 15  
Cys Ser Leu Glu Met Cys Leu Phe Met Ser Asn Ser Ser Leu Pro Thr  
20 25 30  
Lys Pro Asn Arg Lys Thr Arg Phe Gly Asp Arg Cys Leu Leu Met Ala  
35 40 45  
Lys Gln Gln Arg Thr Arg Leu Tyr Ile Leu Arg Arg Cys Val Ser Met  
50 55 60  
Leu Leu Cys Trp His Asp His Ser Ile Ser Asp  
65 70 75

<210> 485  
<211> 55  
<212> PRT  
<213> Arabidopsis thaliana  
<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..55:Ceres Seq. ID 2051908

<400>485  
Met Cys Leu Phe Met Ser Asn Ser Ser Leu Pro Thr Lys Pro Asn Arg  
1 5 10 15  
Lys Thr Arg Phe Gly Asp Arg Cys Leu Leu Met Ala Lys Gln Arg  
20 25 30  
Thr Arg Leu Tyr Ile Leu Arg Arg Cys Val Ser Met Leu Leu Cys Trp  
35 40 45  
His Asp His Ser Ile Ser Asp  
50 55

<210> 486  
<211> 51  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..51:Ceres Seq. ID 2051909

<400>486  
Met Ser Asn Ser Ser Ser Leu Pro Thr Lys Pro Asn Arg Lys Thr Arg Phe  
1 5 10 15  
Gly Asp Arg Cys Leu Leu Met Ala Lys Gln Gln Arg Thr Arg Leu Tyr  
20 25 30  
Ile Leu Arg Arg Cys Val Ser Met Leu Leu Cys Trp His Asp His Ser  
35 40 45  
Ile Ser Asp  
50

<210> 487  
<211> 719  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..719:Ceres Seq. ID 2052403

<400>487  
acaaacatt caaagacata caaaataatt gagtttttt tttaattaga acaaatggg 60  
gttgatggt gtcttcacg tggaggttga ggttaagtct ccggtcgaaa agttctggat 120  
agccctggc gaecgcacga atctctccc caagcttc cctaagcact acaaaacct 180  
ccaagtcta gccggcgatg gaacgcctcc tggctccatt cgcctcatta ctatggaga 240  
agatctcca ctgtgaaaga tatcgccgga gagtatogaa gcagtggatt tggagaacaa 300  
aagcatctg tacagcatca ttggcggaga aatgttgag tactacaaa ccttcaaaag 360  
aacatcacc gttatcccta aggatggtag tagcctctg aaatgctcg gtgagttga 420  
gaagaccgc catggatcgc atgattccca tgcatacag gacttctcg tcaagaact 480  
caaagaga gttagatgc ttcttaaga aactatgcc taacactaga accttaaat 540  
tatataagag gttcgaccg tctctaaag attttctaa ttaagaagtt gaataaag 600  
gaacccttt atgaatccc acgttttatg atttcggagt ttatgcagcc tagtaggca 660  
taagctttt acaaaagcca atttagtoga acattttaa aaaaatgaac ctttgtda 719

<210> 488  
<211> 155  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..155:Ceres Seq. ID 2052404

<400>488  
 Met Gly Leu Ser Gly Val Leu His Val Glu Val Glu Val Lys Ser Pro  
 1 5 10 15  
 Ala Glu Lys Phe Trp Val Ala Leu Gly Asp Gly Ile Asn Leu Phe Pro  
 20 25 30  
 Lys Ala Phe Pro Asn Asp Tyr Lys Thr Ile Gln Val Leu Ala Gly Asp  
 35 40 45  
 Gly Asn Ala Pro Gly Ser Ile Arg Leu Ile Thr Tyr Gly Glu Gly Ser  
 50 55 60  
 Pro Leu Val Lys Ile Ser Ala Glu Arg Ile Glu Ala Val Asp Leu Glu  
 65 70 75 80  
 Asn Lys Ser Met Ser Tyr Ser Ile Ile Gly Gly Glu Met Leu Glu Tyr  
 85 90 95  
 Tyr Lys Thr Phe Lys Gly Thr Ile Thr Val Ile Pro Lys Asp Gly Gly  
 100 105 110  
 Ser Leu Leu Lys Trp Ser Gly Glu Phe Glu Lys Thr Ala His Glu Ile  
 115 120 125  
 Asp Asp Pro His Val Ile Lys Asp Phe Ala Val Lys Asn Phe Lys Glu  
 130 135 140  
 Ile Asp Glu Tyr Leu Leu Lys Gln Thr Ser Ala  
 145 150 155

<210> 489  
 <211> 739  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..739:Ceres Seq. ID 2053545

<400>489  
 attttagtc attgttgacc attcttttccg tatagactac tatctctgat ctcttcgag  
 60  
 ttaagtcag aactagaaa atcagaagc gctctcaatc tcaaaatat ccatggcgc  
 120  
 gattacaga ttctaccaa aagatcagg atatttgtt ctctctctg tcttctactg  
 180  
 ttctctcaac ctctggatgg gtgtcgaagt cgcgaagct cgcgaaggt acacgtccc  
 240  
 gtaccact ctatctgcaa tgaatcaga aacaaagatg gctaaagct tcaactgtgt  
 300  
 tcaagagaga catcaaaact ctttagagat gatgccaatg tatttctaac tgaatgacct  
 360  
 cgtggggatg aagcacctt gtatctgtac tggccttggt ttgctttaca acgttagccg  
 420  
 attctcttac ttaaaaggtt atgctactgg aagctccatg aagcgtctta cgcaggggaa  
 480  
 atacggttct ttgggttgc tagtgtctgat gatctgacc atctcgtttg gtgtcaactt  
 540  
 gactctgct tgaactactc gttttgggg ttaatgattc tctggtttgc tctgaagaata  
 600  
 tagaacaact gctgtgaagc tgcacacaaa acttctgtaa tacittagag ttgtcactt  
 660  
 ttaaaagttt tgaataaact atggtcttcac agaaagtttg aaatttcaaa tccgtagacg  
 720  
 ttaataaaga tttagaatta  
 739

<210> 490  
 <211> 146  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..146:Ceres Seq. ID 2053546

<400>490  
 Met Ala Ala Ile Thr Glu Phe Leu Pro Lys Glu Tyr Gly Tyr Val Val  
 1 5 10 15  
 Leu Val Leu Val Phe Tyr Cys Phe Leu Asn Leu Trp Met Gly Ala Gln  
 20 25 30  
 Val Gly Arg Ala Arg Lys Arg Tyr Asn Val Pro Tyr Pro Thr Leu Tyr  
 35 40 45  
 Ala Ile Glu Ser Glu Asn Lys Asp Ala Lys Leu Phe Asn Cys Val Gln  
 50 55 60  
 Arg Gly His Gln Asn Ser Leu Glu Met Met Pro Met Tyr Phe Ile Leu  
 65 70 75 80  
 Met Ile Leu Gly Gly Met Lys His Pro Cys Ile Cys Thr Gly Leu Gly  
 85 90 95  
 Leu Leu Tyr Asn Val Ser Arg Phe Phe Tyr Phe Lys Gly Tyr Ala Thr  
 100 105 110  
 Gly Asp Pro Met Lys Arg Leu Thr Ile Gly Lys Tyr Gly Phe Leu Gly  
 115 120 125  
 Leu Leu Gly Leu Met Ile Cys Thr Ile Ser Phe Gly Val Thr Leu Ile  
 130 135 140  
 Leu Ala  
 145

<210> 491  
 <211> 118  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..118:Ceres Seq. ID 2053547

<400>491  
 Met Gly Ala Gln Val Gly Arg Ala Arg Lys Arg Tyr Asn Val Pro Tyr  
 1 5 10 15  
 Pro Thr Leu Tyr Ala Ile Glu Ser Glu Asn Lys Asp Ala Lys Leu Phe  
 20 25 30  
 Asn Cys Val Gln Arg Gly His Gln Asn Ser Leu Glu Met Met Pro Met  
 35 40 45  
 Tyr Phe Ile Leu Met Ile Leu Gly Gly Met Lys His Pro Cys Ile Cys  
 50 55 60  
 Thr Gly Leu Gly Leu Leu Tyr Asn Val Ser Arg Phe Phe Tyr Phe Lys  
 65 70 75 80  
 Gly Tyr Ala Thr Gly Asp Pro Met Lys Arg Leu Thr Ile Gly Lys Tyr  
 85 90 95  
 Gly Phe Leu Gly Leu Leu Gly Leu Met Ile Cys Thr Ile Ser Phe Gly  
 100 105 110  
 Val Thr Leu Ile Leu Ala  
 115

<210> 492  
 <211> 74  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>

<223> any n or Xaa = unknown

<223> LOCATION: 1..74:Ceres Seq. ID 2053548

<400>492  
 Met Met Pro Met Tyr Phe Ile Leu Met Ile Leu Gly Gly Met Lys His  
 1 5 10 15  
 Pro Cys Ile Cys Thr Gly Leu Gly Leu Tyr Asn Val Ser Arg Phe  
 20 25 30  
 Phe Tyr Phe Lys Gly Tyr Ala Thr Gly Asp Pro Met Lys Arg Leu Thr  
 35 40 45  
 Ile Gly Lys Tyr Gly Phe Leu Gly Leu Gly Leu Met Ile Cys Thr  
 50 55 60  
 Ile Ser Phe Gly Val Thr Leu Ile Leu Ala  
 65 70

<210> 493  
 <211> 681  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..681:Ceres Seq. ID 2053884

<400>493  
 ccattcaaaa aaaaacaac aaaaattat attcaagsa aaaaagaaa atgaatttc  
 60  
 atctccgata agttaagaac actctcaagc tcaacaccag agagccaga caacaacaag  
 120  
 ccagctgaag gaaccgaac actacaaga ccagctacca agccgagct catgccaagt  
 180  
 gccaaagt tagctgaagc tctcaagcc gcagctcgt aagaatcaga caaactgac  
 240  
 aaggttaag tgcgcgagc ctctgtgat atttagacg ctgcgcgaa atacggtaag  
 300  
 ttcatgaaa agagtgcac tggctcgtac ctgcgaagg ctgagaagta tctcaagac  
 360  
 tacagctcgt cacaactcac cggctctggt ggctcctc ctcgcagag tcagctgag  
 420  
 ccaagaagc agctcagcc ggcgcgaag aagacgaig aagagctgg tggctggctt  
 480  
 ggaagtatg ccaagtgcg tcaagtttc ttgaagtat ttgacttta atgtgttc  
 540  
 acatttttg taataataa ttaataact agtactggtt gtagctgatt tatgtgctt  
 600  
 cgttatggt tatgggagt gacgagtgag tgaataact tctggtgatc atgaatcaa  
 660  
 tcatctcttg ttgtgattat g  
 681

<210> 494  
 <211> 172  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..172:Ceres Seq. ID 2053885

<400>494  
 Pro Ser Lys Lys Asn Lys Thr Lys Asn Tyr Ile Gln Gly Lys Lys Lys  
 1 5 10 15  
 Lys Met Asn Phe Ile Ser Asp Gln Val Lys Lys Lys Ser Ser Thr  
 20 25 30  
 Pro Glu Glu Pro Asp His Asn Lys Pro Val Glu Gly Thr Glu Thr Ala  
 35 40 45  
 Thr Arg Pro Ala Thr Asn Ala Glu Leu Met Ala Ser Ala Lys Val Val  
 50 55 60  
 Ala Glu Ala Ala Gln Ala Ala Arg Asn Glu Ser Asp Lys Leu Asp  
 65 70 75  
 Lys Gly Lys Val Ala Gly Ala Ser Ala Asp Ile Leu Asp Ala Ala Glu  
 80 85 90 95  
 Lys Tyr Gly Lys Phe Asp Glu Lys Ser Ser Thr Gly Gln Tyr Leu Asp  
 100 105 110

Lys Ala Glu Lys Tyr Tyr Asn Asp Tyr Glu Ser Ser His Ser Thr Gly  
 115 120 125  
 Ala Gly Gly Pro Pro Thr Ser Gln Ala Glu Pro Ala Ser Gln  
 130 135 140  
 Pro Glu Pro Ala Ala Lys Lys Asp Asp Glu Glu Ser Gly Gly Lys  
 145 150 155 160  
 Gly Gly Tyr Ala Lys Met Ala Gln Gly Phe Leu Lys  
 165 170

<210> 495  
 <211> 155  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..155:Ceres Seq. ID 2053886

<400>495  
 Met Asn Phe Ile Ser Asp Gln Val Lys Lys Leu Ser Ser Thr Thr Pro  
 1 5 10 15  
 Glu Glu Pro Asp His Asn Lys Pro Val Glu Gly Thr Glu Thr Ala Thr  
 20 25 30  
 Arg Pro Ala Thr Asn Ala Glu Leu Met Ala Ser Ala Lys Val Val Ala  
 35 40 45  
 Glu Ala Ala Gln Ala Ala Arg Asn Glu Ser Asp Lys Leu Asp Lys  
 50 55 60  
 Gly Lys Val Ala Gly Ala Ser Ala Asp Ile Leu Asp Ala Ala Glu Lys  
 65 70 75 80  
 Tyr Gly Lys Phe Asp Glu Lys Ser Ser Thr Gly Gln Tyr Leu Asp Lys  
 85 90 95  
 Ala Glu Lys Tyr Leu Asn Asp Tyr Glu Ser Ser His Ser Thr Gly Ala  
 100 105 110  
 Gly Gly Pro Pro Pro Thr Thr Ser Gln Ala Glu Pro Ala Ser Gln Pro  
 115 120 125  
 Glu Pro Ala Ala Lys Lys Asp Asp Glu Glu Ser Gly Gly Lys Leu Gly  
 130 135 140  
 Gly Tyr Ala Lys Met Ala Gln Gly Phe Leu Lys  
 145 150 155

<210> 496  
 <211> 115  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..115:Ceres Seq. ID 2053887

<400>496  
 Met Ala Ser Ala Lys Val Val Ala Glu Ala Ala Gln Ala Ala Arg  
 1 5 10 15  
 Asn Glu Ser Asp Lys Leu Asp Lys Gly Lys Val Ala Gly Ala Ser Ala  
 20 25 30  
 Asp Ile Leu Asp Ala Ala Glu Lys Tyr Gly Lys Phe Asp Glu Lys Ser  
 35 40 45  
 Ser Thr Gly Gln Tyr Leu Asp Lys Ala Glu Lys Tyr Leu Asn Asp Tyr  
 50 55 60  
 Glu Ser Ser His Ser Thr Gly Ala Gly Pro Pro Pro Thr Thr Ser  
 160

65 Gln Ala Glu Pro Ala Ser Gln Pro Glu Pro Ala Ala Lys Lys Asp Asp 80  
85 90  
Glu Glu Ser Gly Gly Lys Lys Gly Tyr Ala Lys Met Ala Gln Gly 110  
100  
Phe Leu Lys 115

&lt;210&gt; 497

&lt;211&gt; 797

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..797:Ceres Seq. ID 2053908

&lt;400&gt;497

ttcttcaaa atccataana agagagagag ataataaag agaaaactaa agaagctaga 60  
agatgagaa aegtaaatg catgacaagg ctgaccacgg cggctccggt gggtggcca 120  
cgggaagtg ggaggaagt agctcagaa tccgaactgc cgaacaaatg ctccggttag 180  
ctccgttgg ccttgggtt gcgaacatg tctgcatgct taagaaactt gagactaatg 240  
agctcgttc aattctttac tccaatctca cagcctttag gtacttggg cagcaaatg 300  
gaatagtgc aggcactct cttctatcag cagccattgc agcgatgctt cgttcttct 360  
cgacaatgc tctgtttgg acctctttt gctcgaacca gcttctgacc tacttggttc 420  
ttgtgtcgg agctgatca gcagaggttc tatactggc ttacaatgga gactcagcca 480  
ttacttggg cgtatcagt agctcattat gcggtttctg tctatagacc actgctctg 540  
ttataatcac attcttggg gttttttct acatggttct cttctaatc tctcttata 600  
agctcttgc tctcttggat cctcttcca ttgttgactc cgcacaagaat ctgaagtgc 660  
ctgtcttcgg aattgatgc tctcattgac tcttcaatgc cttctatggtt gctatgttc 720  
aaatgtgtg tgtttttac ttgtgtcagg agaaagaatg cttatgtgtt tgcctctct 780  
tctatgcttt tcttctt 797

&lt;210&gt; 498

&lt;211&gt; 204

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..204:Ceres Seq. ID 2053909

&lt;400&gt;498

Met Glu Lys Ser Asn Asp His Asp Lys Ala Ser His Gly Gly Ser Gly 1  
5 10 15  
Gly Gly Ala Thr Glu Lys Trp Glu Glu Thr Ser Leu Gly Ile Arg Thr 20  
25 30  
Ala Glu Thr Met Leu Arg Leu Ala Pro Val Gly Leu Cys Val Ala Ala 35  
35 40 45  
Leu Val Val Met Leu Lys Asp Ser Glu Thr Asn Glu Phe Gly Ser Ile 50  
55 60  
Ser Tyr Ser Asn Leu Thr Ala Phe Arg Tyr Leu Val His Ala Asn Gly 65  
70 75 80  
Ile Cys Ala Gly Tyr Ser Leu Leu Ser Ala Ala Ile Ala Ala Met Pro 85  
90 95  
Arg Ser Ser Ser Thr Met Pro Arg Val Trp Thr Phe Phe Cys Leu Asp 100  
105 110  
Gln Leu Leu Thr Tyr Leu Val Leu Ala Ala Gly Ala Val Ser Ala Glu 115  
120 125

Val Leu Tyr Leu Ala Tyr Asn Gly Asp Ser Ala Ile Thr Trp Ser Asp 130  
135 140  
Ala Cys Ser Ser Tyr Gly Gly Phe Cys His Arg Ala Thr Ala Ser Val 145  
150 155  
Ile Ile Thr Phe Phe Val Val Cys Phe Tyr Ile Val Leu Ser Leu Ile 160  
165 170 175  
Ser Ser Tyr Lys Leu Phe Thr Arg Phe Asp Pro Ser Ile Val Asp 180  
185 190  
Ser Ala Lys Asn Leu Glu Val Ala Val Phe Gly Ser 195  
200

&lt;210&gt; 499

&lt;211&gt; 169

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..169:Ceres Seq. ID 2053910

&lt;400&gt;499

Met Leu Arg Leu Ala Pro Val Gly Leu Cys Val Ala Ala Leu Val Val 1  
5 10 15  
Met Leu Lys Asp Ser Glu Thr Asn Glu Phe Gly Ser Ile Ser Tyr Ser 20  
25 30  
Asn Leu Thr Ala Phe Arg Tyr Leu Val His Ala Asn Gly Ile Cys Ala 35  
40 45  
Gly Tyr Ser Leu Leu Ser Ala Ala Ile Ala Ala Met Pro Arg Ser Ser 50  
55 60  
Ser Thr Met Pro Arg Val Trp Thr Phe Phe Cys Leu Asp Gln Leu Leu 65  
70 75 80  
Thr Tyr Leu Val Leu Ala Ala Gly Ala Val Ser Ala Glu Val Leu Tyr 85  
90 95  
Leu Ala Tyr Asn Gly Asp Ser Ala Ile Thr Trp Ser Asp Ala Cys Ser 100  
105 110  
Ser Tyr Gly Gly Phe Cys His Arg Ala Thr Ala Ser Val Ile Ile Thr 115  
120 125  
Phe Phe Val Val Cys Phe Tyr Ile Val Leu Ser Leu Ile Ser Ser Tyr 130  
135 140  
Lys Leu Phe Thr Arg Phe Asp Pro Ser Ile Val Asp Ser Ala Lys 145  
150 155  
Asn Leu Glu Val Ala Val Phe Gly Ser 160  
165

&lt;210&gt; 500

&lt;211&gt; 153

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..153:Ceres Seq. ID 2053911

&lt;400&gt;500

Met Leu Lys Asp Ser Glu Thr Asn Glu Phe Gly Ser Ile Ser Tyr Ser 1  
5 10 15  
Asn Leu Thr Ala Phe Arg Tyr Leu Val His Ala Asn Gly Ile Cys Ala 20  
25 30  
Gly Tyr Ser Leu Leu Ser Ala Ala Ile Ala Ala Met Pro Arg Ser Ser 115  
120

35 Ser Thr Met Pro Arg Val Thr Thr Phe Cys Leu Asp Gln Leu Leu  
50 55 60  
Thr Tyr Leu Val Leu Ala Ala Gly Ala Val Ser Ala Glu Val Leu Tyr  
65 70 75 80  
Leu Ala Tyr Asn Gly Asp Ser Ala Ile Thr Trp Ser Asp Ala Cys Ser  
85 90 95  
Ser Tyr Gly Gly Phe Cys His Arg Ala Thr Ala Ser Val Ile Ile Thr  
100 105 110  
Phe Phe Val Val Cys Phe Tyr Ile Val Leu Ser Ile Ser Ser Tyr  
115 120 125  
Lys Leu Phe Thr Arg Phe Asp Pro Pro Ser Ile Val Asp Ser Ala Lys  
130 135 140  
Asn Leu Glu Val Ala Val Phe Gly Ser  
145 150

<210> 501  
<211> 652  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..652:Ceres Seq. ID 2056123

<400>501  
atttttttctctctctctcctccctaagc aaaaactaaaa caagctatgg ctggtatgct  
tccggagtt gaggtagca gggggggggcgt ctccacgggt ggtgtctctc cgtatgaac  
ctcgaacaca gctctgtggt cggctcgggc gggacacgtc tggacacggc gaccatggt  
ctctctttac actaccaatc atgagagcca cgaaggcccat gctctcttct cggagagaag  
tgttagaat aatcttatg gagaagacaa cgtatgagaa cttagcggag cagcaaga  
ggcaacagg aggtttaaca agcggctgag aatccacca cgtacaaggc aaaaatggtaa  
agacaagggt aataaatgtt agcaaggaaa gggtaaacct ctccgggact taccgaccga  
ggtgtcggt ttaaaaga ggcagagaag gttgatgaa tggttcaagc ggcgagttag  
ggacaacaa gatttgcta tatgtctaga cggtttcaag aggggtgaga ccttggtaca  
ctccatagt gccataagt ttaactcat atgcttatgt ccttggttag acacaatgt  
ttattgcca tatttgaga ctgatatttg gaattaaatg ttatatttt ga

<210> 502  
<211> 196  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..196:Ceres Seq. ID 2056124

<400>502  
Met Ala Gly Met Leu Pro Gly Val Glu Cys Ala Arg Arg Arg Arg Phe  
1 5 10 15  
His Gly Gly Ala Pro Pro Ile Glu Ser Ser Asn Thr Ala Ser Val Ala  
20 25 30  
Ala Ala Ala Gly His Val Thr Thr Arg Arg Pro Ser Phe Ser Leu Tyr  
35 40 45  
Thr Thr Asn His Glu Ser His Gln Ala His Val Ser Phe Ser Glu Arg  
50 55 60  
Ser Val Arg Asn Lys Ser Tyr Gly Glu Asp Asn Asp Glu Lys Leu Asp  
65 70 75 80  
Gly Ala Ala Lys Glu Ala Lys Gln Arg Leu Asn Lys Arg Leu Arg Ile

85 90 95  
Pro Pro Arg Thr Arg Gln Asn Gly Lys Asp Lys Gly Asn Lys Leu Glu  
100 105 110  
Gln Gly Lys Gly Lys Pro Leu Gly Asp Leu Pro Thr Thr Gln Val Val Gly  
115 120 125  
Leu Lys Lys Ser Arg Gly Arg Leu Met Glu Trp Phe Lys Arg Arg Val  
130 135 140  
Arg Glu Gln Gln Asp Cys Ala Ile Cys Leu Asp Arg Phe Lys Lys Gly  
145 150 155  
Glu Thr Leu Val His Leu Pro Cys Ala His Lys Phe His Ser Ile Cys  
165 170 175  
Leu Leu Pro Trp Leu Asp Thr Asn Val Tyr Cys Pro Tyr Cys Arg Thr  
180 185 190  
Asp Ile Trp Asn  
195

<210> 503  
<211> 193  
<212> PRT  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

<223> LOCATION: 1..193:Ceres Seq. ID 2056125

<400>503  
Met Leu Pro Gly Val Glu Cys Ala Arg Arg Arg Arg Phe His Gly Gly  
1 5 10 15  
Ala Pro Pro Ile Glu Ser Ser Asn Thr Ala Ser Val Ala Ala Ala  
20 25 30  
Gly His Val Trp Thr Arg Arg Pro Ser Phe Ser Leu Tyr Thr Thr Asn  
35 40 45  
His Glu Ser His Gln Ala His Val Ser Phe Ser Glu Arg Ser Val Arg  
50 55 60  
Asn Lys Ser Tyr Gly Glu Asp Asn Asp Glu Lys Leu Asp Gly Ala Ala  
65 70 75 80  
Lys Glu Ala Lys Gln Arg Leu Asn Lys Arg Leu Arg Ile Pro Pro Arg  
85 90 95  
Thr Arg Gln Asn Gly Lys Asp Lys Gly Asn Lys Leu Glu Gln Gly Lys  
100 105 110  
Gly Lys Pro Leu Gly Asp Leu Pro Thr Glu Val Val Gly Leu Lys Lys  
115 120 125  
Ser Arg Gly Arg Leu Met Glu Trp Phe Lys Arg Arg Val Arg Glu Gln  
130 135 140  
Gln Asp Cys Ala Ile Cys Leu Asp Arg Phe Lys Lys Gly Glu Thr Leu  
145 150 155  
Val His Leu Pro Cys Ala His Lys Phe His Ser Ile Cys Leu Leu Pro  
165 170 175  
Trp Leu Asp Thr Asn Val Tyr Cys Pro Tyr Cys Arg Thr Asp Ile Trp  
180 185 190  
Asn

<210> 504  
<211> 2619  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<223> any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..2619:Ceres Seq. ID 2056245

&lt;400&gt;504

atggagcaag agaaagatt ggaaccaca ctaggcgatg cttgtcagg atcaatggtt 60  
 caatccct cactgaatt acggtttt tacttcgtc aaggcaaac agagacgct 120  
 cagccctc cagtattca cgcgcgcgc ttcacacct ttatctctg tgcgtcgtc 180  
 tccgtgaat tccgcgcga cgcgaacaa gagaagtt ttgctaaat taccgtttg 240  
 ccaattccg gaacgactt ggaatagaa acgacgcgc ttttggtct aactctcct 300  
 tcttcgag gtaacgtaa cgttaagaa aaacgcgct ctttcgtaa aacgttaag 360  
 cagctcagc caataacgg cgttggttc tccgtccac gttatgccc cgagacgatt 420  
 ttcccgccg ttattatc ggcggagca cgggttcana ccgtgatgc taagaacatc 480  
 caaggcga ctygaatt cggcatatt tacagagaa caccctgcgc taatcccta 540  
 acacacggtt ggaacatt cgttaacag agaaactaa tgcgcgaga ctaaatcgtc 600  
 ttctccgt cgaatccgg taccctcgc ggcgaatcc cccgcgcgt acggcggtt 660  
 ctggatcta acgagggatc cgaatcct tccctggat tctccggtt cctccgtgac 720  
 gacgagctca caacacacac atcgaagcta atgacatga aacgaacgc aaacacgac 780  
 gaaacgcgc cgttacacag gagggttaga tagaagcag tagcgaagc ggtggcggt 840  
 gcagcgtg gacaacgct ttgagttgt tattatccac gcgtatgac accgagttt 900  
 tgcgtaaag cagctgagt tagatcgca atcgagataa gatggttag tggatcgtt 960  
 tttaaatg cgttgaac agagattct ttgagataa gttggttat ggtactcgt 1020  
 tcccgcttc aagtcgca tccaatcgt tgcctaatt caccatggc tctctcag 1080  
 gtgcttgg acgaacccga ttgttaca acgttaagc ggttgatcc ggttgatc 1140  
 gaattgat cgaacatcg taacatcat ttatctcat tctctcag gaagaatt 1200  
 aggattccgc agcaattga gtttcattc caggtacta aattccgat ttctcccg 1260  
 ggaatgcga acaatggcg tggcagtc atgtgtatc tgtcaacga caacaata 1320  
 gctcccgag gaatacagg agccaggca gtcacnac tcttcggacc accatctccg 1380  
 tctttgtt ctgactcaca ttttagtag tacacgcta acaacagtt acattctcg 1440  
 gcaatgtc talcagttt caaccaggg catcctatt atcagctag ggaatgag 1500  
 aaagaata acattctcg tcttaact atggggatc ctgctatgt tccggataag 1560  
 aagaatctg ttgtrcgt taagacat cagttcgt tcttcggtca accgattta 1620  
 accgaacgc aagtataga ccgaacagg ttttggag aagagcgga agcggaggag 1680  
 gagaagtt tagtgctcg tgggttaaca tgaattata gtttgaag acttgagacg 1740  
 ggcattga agtttcat ggaactgag gatgtggac gaaactcga tctctcgtt 1800  
 atggtcgtt accaagaat gtaacgaaa ttgctgaga gtttcatat agaagagagg 1860  
 tccgattgt tgcactatg tgttacgg gatgcaaat gtttatcaa acgtattgga 1920  
 gacgaact tcaagatt catgaagca actaaacgc taacataca gatggatat 1980  
 ggtggcgca acgtgagaa cattttagt tccctcaa taatatttt tctctttt 2040  
 ttcaaaaat taattttt ttgtttct tccatggtg gatctatc tcatatcat 2100  
 atatatcat cagattat atatatat atatatat atataatga taataaaaa 2160  
 aaacttatt tgtcatct ttaagtta aaacagta atctattat atcatctag 2220  
 atgaagaac taattcaact cctgtcaat gttatgata tcttcccat tcttgcta 2280  
 ggaagtgtt caaatgagg gctagaaa ccaaaaaa aatgataa gattctaaag 2340  
 caatccact gttgtcgt agagtgcac tctatgtg tgaagaagg ggggaagga 2400  
 gccggtact gcttcttc taaagagt taccgtatt atcatgcc ttacgattt 2460  
 tattattac aataataga gatttatg aaatagtaa aatagtaat cgggtgatta 2520  
 tagcctaac atgatatag ttgtctcat caataaact atatagacta aaagtaataa 2580  
 taataatgac aatcataact cagaagttt ttgttagc 2619

&lt;210&gt; 505

&lt;211&gt; 817

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or kaa = unknown

&lt;223&gt; LOCATION: 1..817:Ceres Seq. ID 2056246

&lt;400&gt;505

Met Glu Gln Glu Lys Ser Leu Asp Pro Gln Leu Trp His Ala Cys Ala

1 10 15  
 Gly Ser Met Val Gln Ile Pro Ser Leu Asn Ser Thr Val Phe Tyr Phe  
 20 25 30  
 Ala Gln Gly His Thr Glu His Ala Ala Pro Pro Asp Phe His Ala  
 35 40 45  
 Pro Arg Val Pro Pro Leu Ile Leu Cys Arg Val Val Ser Val Lys Phe  
 50 55 60  
 Leu Ala Asp Ala Glu Thr Asp Glu Val Phe Ala Lys Ile Thr Leu Leu  
 65 70 75 80  
 Pro Leu Pro Gly Asn Asp Leu Asp Leu Glu Asn Asp Ala Val Leu Gly  
 85 90 95  
 Leu Thr Pro Pro Ser Ser Asp Gly Asn Gly Asn Gly Lys Glu Lys Pro  
 100 105 110  
 Ala Ser Phe Ala Lys Thr Leu Thr Gln Ser Asp Ala Asn Asn Gly Gly  
 115 120 125  
 Gly Phe Ser Val Pro Arg Tyr Cys Ala Glu Thr Ile Phe Pro Arg Leu  
 130 135 140  
 Asp Tyr Ser Ala Glu Pro Pro Val Gln Thr Val Ile Ala Lys Asp Ile  
 145 150 155 160  
 His Gly Glu Thr Trp Lys Phe Arg His Ile Tyr Arg Gly Thr Pro Arg  
 165 170 175  
 Arg His Leu Leu Thr Thr Gly Trp Ser Thr Phe Val Asn Gln Lys Lys  
 180 185 190  
 Leu Ile Ala Gly Asp Ser Ile Val Phe Leu Arg Ser Glu Ser Gly Asp  
 195 200 205  
 Leu Cys Val Gly Ile Arg Arg Ala Lys Arg Gly Leu Gly Ser Asn  
 210 215 220  
 Ala Gly Ser Asp Asn Pro Tyr Pro Gly Phe Ser Gly Phe Leu Arg Asp  
 225 230 235 240  
 Asp Glu Ser Thr Thr Thr Ser Lys Leu Met Met Met Lys Arg Asn  
 245 250 255  
 Gly Asn Asn Asp Gly Asn Ala Ala Thr Gly Arg Val Arg Val Glu  
 260 265 270  
 Ala Val Ala Glu Ala Val Ala Arg Ala Ala Cys Gly Gln Ala Phe Glu  
 275 280 285  
 Val Val Tyr Tyr Pro Arg Ala Ser Thr Pro Glu Phe Cys Val Lys Ala  
 290 295 300  
 Ala Asp Val Arg Ser Ala Met Arg Ile Arg Trp Cys Ser Gly Met Arg  
 305 310 315  
 Phe Lys Met Ala Phe Glu Thr Glu Asp Ser Ser Arg Ile Ser Trp Phe  
 320 325 330 335  
 Met Gly Thr Val Ser Ala Val Gln Val Ala Asp Pro Ile Arg Trp Pro  
 340 345 350  
 Asn Ser Pro Trp Arg Leu Leu Gln Val Ala Trp Asp Glu Pro Asp Leu  
 355 360 365  
 Leu Gln Asn Val Lys Arg Val Ser Pro Trp Leu Val Glu Leu Val Ser  
 370 375 380  
 Asn Met Pro Thr Ile His Leu Ser Pro Phe Ser Pro Arg Lys Lys Ile  
 385 390 395 400  
 Arg Ile Pro Gln Pro Phe Glu Phe Pro Phe His Gly Thr Lys Phe Pro  
 405 410 415  
 Ile Phe Ser Pro Gly Phe Ala Asn Asn Gly Gly Gly Ser Met Cys  
 420 425 430  
 Tyr Leu Ser Asn Asp Asn Asn Ala Pro Ala Gly Ile Gln Gly Ala  
 435 440 445  
 Arg Gln Ala Gln Gln Leu Phe Gly Ser Pro Ser Pro Ser Leu Leu Ser  
 450 455 460  
 Asp Leu Asn Leu Ser Ser Tyr Thr Gly Asn Asn Lys Leu His Ser Pro  
 465 470 475 480  
 Ala Met Phe Leu Ser Ser Phe Asn Pro Arg His His Tyr Gln Ala  
 485 490 495

Arg Asp Ser Glu Asn Ser Asn Asn Ile Ser Cys Ser Leu Thr Met Gly 500  
 Asn Pro Ala Met Val Gln Asp Lys Lys Ser Val Gly Ser Val Lys 510  
 Thr His Gln Phe Val Leu Phe Gly Gln Pro Ile Leu Thr Glu Gln Gln 525  
 Val Met Asn Arg Lys Arg Phe Leu Glu Gln Ala Glu Ala Glu Glu 540  
 545 Gly Lys Glu Val Ala Arg Gly Leu Thr Trp Asn Tyr Ser Leu Gln 560  
 565 Gly Leu Glu Thr Gly His Cys Lys Val Phe Met Glu Ser Glu Asp Val 575  
 Gly Arg Thr Leu Asp Leu Ser Val Ile Gly Ser Tyr Gln Glu Leu Tyr 605  
 Arg Lys Leu Ala Glu Met Phe His Ile Glu Glu Arg Ser Asp Leu Leu 620  
 Thr His Val Val Tyr Arg Asp Ala Asn Gly Val Ile Lys Arg Ile Gly 640  
 625 Asp Glu Pro Phe Ser Asp Phe Met Lys Ala Thr Lys Arg Leu Thr Ile 655  
 Lys Met Asp Ile Gly Gly Asp Asn Val Arg Asn Ile Phe Ser Phe Leu 670  
 Gln Ile Ile Phe Phe Phe Phe Lys Lys Lys Leu Ile Phe Phe Leu 685  
 Phe Phe Ser Asp Gly Gly Tyr Tyr Thr His Ile His Ile Tyr Ser Ser 700  
 690 Asp Tyr Ile Tyr Ile Tyr Ile Tyr Ile Ser Asn Ile Lys Lys 715  
 705 Lys Leu Ile Leu Phe Ile Leu Leu Ser Leu Lys Thr Val Asn Ser Phe 735  
 725 Ile Ser Ser Met Met Lys Lys Leu Ile Gln Leu Ser Phe Thr Val Met 750  
 740 Ile Ile Phe Thr Ile Leu Val Leu Gly Val Val Ala Asn Glu Gly Leu 765  
 755 Gly Lys Pro Lys Lys Gln Cys Asn Glu Ile Leu Lys Gln Ser Asn Cys 780  
 770 Val Ala Ala Glu Cys Asp Ser Met Cys Val Lys Lys Arg Gly Lys Gly 800  
 785 Ala Gly Tyr Cys Ser Pro Ser Lys Lys Cys Tyr Tyr Tyr His Cys 815  
 805 Pro

&lt;210&gt; 506

&lt;211&gt; 799

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..799;Ceres Seq. ID 2056247

&lt;400&gt;506

Met Val Gln Ile Pro Ser Leu Asn Ser Thr Val Phe Tyr Phe Ala Gln 10  
 1 Gly His Thr Glu His Ala His Ala Pro Pro Asp Phe His Ala Pro Arg 20  
 Val Pro Pro Leu Ile Leu Cys Arg Val Val Ser Val Lys Phe Leu Ala 35  
 Asp Ala Glu Thr Asp Glu Val Phe Ala Lys Ile Thr Leu Leu Pro Leu 40

Pro Gly Asn Asp Leu Asp Leu Glu Asn Asp Ala Val Leu Gly Leu Thr 60  
 65 Pro Pro Ser Ser Asp Gly Asn Gly Asn Gly Lys Glu Lys Pro Ala Ser 75  
 85 Phe Ala Lys Thr Leu Thr Gln Ser Asp Ala Asn Asn Gly Gly Phe 95  
 100 Ser Val Pro Arg Tyr Cys Ala Glu Thr Ile Phe Pro Arg Leu Asp Tyr 110  
 115 Ser Ala Glu Pro Pro Val Gln Thr Val Ile Ala Lys Asp Ile His Gly 125  
 130 Glu Thr Trp Lys Phe Arg His Ile Tyr Arg Gly Thr Pro Arg Arg His 140  
 145 Leu Leu Thr Thr Gly Trp Ser Thr Phe Val Asn Gln Lys Lys Leu Ile 155  
 Ala Gly Asp Ser Ile Val Phe Leu Arg Ser Glu Ser Gly Asp Leu Cys 170  
 Val Gly Ile Arg Arg Ala Lys Arg Gly Gly Leu Gly Ser Asn Ala Gly 185  
 Ser Asp Asn Pro Tyr Pro Gly Phe Ser Gly Phe Leu Arg Asp Asp Glu 200  
 210 Ser Thr Thr Thr Ser Lys Leu Met Met Met Lys Arg Asn Gly Asn 220  
 225 Asn Asp Gly Asn Ala Ala Thr Gly Arg Val Arg Val Glu Ala Val 235  
 Ala Glu Ala Val Ala Arg Ala Ala Cys Gly Gln Ala Phe Glu Val Val 250  
 Tyr Tyr Pro Arg Ala Ser Thr Pro Glu Phe Cys Val Lys Ala Ala Asp 265  
 Val Arg Ser Ala Met Arg Ile Arg Trp Cys Ser Gly Met Arg Phe Lys 285  
 Met Ala Phe Glu Thr Glu Asp Ser Ser Arg Ile Ser Trp Phe Met Gly 300  
 305 Thr Val Ser Ala Val Gln Val Ala Asp Pro Ile Arg Trp Pro Asn Ser 315  
 Pro Trp Arg Leu Leu Gln Val Ala Trp Asp Glu Pro Asp Leu Leu Gln 330  
 340 Asn Val Lys Arg Val Ser Pro Trp Leu Val Glu Leu Val Ser Asn Met 345  
 355 Pro Thr Ile His Leu Ser Pro Phe Ser Pro Arg Lys Lys Ile Arg Ile 365  
 370 Pro Gln Pro Phe Phe Glu Phe Pro Phe His Gly Thr Lys Phe Pro Ile Phe 380  
 385 Ser Pro Gly Phe Ala Asn Asn Gly Gly Glu Ser Met Cys Tyr Leu 395  
 405 Ser Asn Asp Asn Asn Ala Pro Ala Gly Ile Gln Gly Ala Arg Gln 410  
 Ala Gln Gln Leu Phe Gly Ser Pro Ser Pro Ser Leu Leu Ser Asp Leu 425  
 435 Asn Leu Ser Ser Tyr Thr Gly Asn Asn Lys Leu His Ser Pro Ala Met 445  
 450 Phe Leu Ser Ser Phe Asn Pro Arg His His Tyr Gln Ala Arg Asp 460  
 465 Ser Glu Asn Ser Asn Asn Ile Ser Cys Ser Leu Thr Met Gly Asn Pro 475  
 485 Ala Met Val Gln Asp Lys Lys Lys Ser Val Gly Ser Val Lys Thr His 495  
 500 Gln Phe Val Leu Phe Gly Gln Pro Ile Leu Thr Glu Gln Val Met 510  
 515 Asn Arg Lys Arg Phe Leu Glu Glu Ala Glu Ala Glu Glu Lys 525  
 530



Gly Leu Val Ala Arg Gly Leu Thr Trp Asn Tyr Ser Leu Gln Gly Leu  
 545 550 555 560  
 Glu Thr Gly His Cys Lys Val Phe Met Glu Ser Glu Asp Val Gly Arg  
 565 570 575  
 Thr Leu Asp Leu Ser Val Ile Gly Ser Tyr Gln Glu Leu Tyr Arg Lys  
 580 585 590 595  
 Leu Ala Glu Met Phe His Ile Glu Glu Arg Ser Asp Leu Leu Thr His  
 595 600 605  
 Val Val Tyr Arg Asp Ala Asn Gly Val Ile Lys Arg Ile Gly Asp Glu  
 610 615 620  
 Pro Phe Ser Asp Phe Met Lys Ala Thr Lys Arg Leu Thr Ile Lys Met  
 625 630 635 640  
 Asp Ile Gly Gly Asp Asn Val Arg Asn Ile Phe Ser Phe Leu Gln Ile  
 645 650 655  
 Ile Phe Phe Phe Phe Phe Lys Lys Leu Ile Phe Phe Leu Phe Phe  
 660 665 670  
 Ser Asp Gly Gly Tyr Tyr Thr His Ile His Ile Tyr Ser Ser Asp Tyr  
 675 680 685  
 Ile Tyr Ile Tyr Ile Tyr Ile Ser Asn Ile Lys Lys Lys Leu  
 690 695 700  
 Ile Leu Phe Ile Leu Leu Ser Leu Lys Thr Val Asn Ser Phe Ile Ser  
 705 710 715 720  
 Ser Met Met Lys Lys Leu Ile Gln Leu Ser Phe Thr Val Met Ile Ile  
 725 730 735  
 Phe Thr Ile Leu Val Leu Gly Val Val Ala Asn Glu Gly Leu Gly Lys  
 740 745 750  
 Pro Lys Lys Gln Cys Asn Glu Ile Leu Lys Gln Ser Asn Cys Val Ala  
 755 760 765  
 Ala Glu Cys Asp Ser Met Cys Val Lys Lys Arg Gly Lys Gly Ala Gly  
 770 775 780  
 Tyr Cys Ser Pro Ser Lys Lys Cys Tyr Cys Tyr Tyr His Cys Pro  
 785 790 795

&lt;210&gt; 507

&lt;211&gt; 567

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..567:Ceres Seq. ID 2056248

&lt;400&gt;507

Met Met Met Lys Arg Asn Gly Asn Asn Asp Gly Asn Ala Ala Thr  
 1 5 10 15  
 Gly Arg Val Arg Val Glu Ala Val Ala Glu Ala Val Ala Arg Ala Ala  
 20 25 30  
 Cys Gly Gln Ala Phe Glu Val Val Tyr Tyr Pro Arg Ala Ser Thr Pro  
 35 40 45  
 Glu Phe Cys Val Lys Ala Ala Asp Val Arg Ser Ala Met Arg Ile Arg  
 50 55 60  
 Trp Cys Ser Gly Met Arg Phe Lys Met Ala Phe Glu Thr Glu Asp Ser  
 65 70 75 80  
 Ser Arg Ile Ser Trp Phe Met Gly Thr Val Ser Ala Val Gln Val Ala  
 85 90 95  
 Asp Pro Ile Arg Trp Pro Asn Ser Pro Trp Arg Leu Leu Gln Val Ala  
 100 105 110  
 Trp Asp Glu Pro Asp Leu Leu Gln Asn Val Lys Arg Val Ser Pro Trp  
 115 120 125  
 Leu Val Glu Leu Val Ser Asn Met Pro Thr Ile His Leu Ser Pro Phe

130 135 140  
 Ser Pro Arg Lys Lys Ile Arg Ile Pro Gln Pro Phe Glu Phe Pro Phe  
 145 150 155 160  
 His Gly Thr Lys Lys Phe Pro Ile Phe Ser Pro Gly Phe Ala Asn Asn Gly  
 165 170 175  
 Gly Gly Glu Ser Met Cys Tyr Leu Ser Asn Asp Asn Asn Ala Pro  
 180 185 190  
 Ala Gly Ile Gln Gly Ala Arg Gln Ala Gln Leu Phe Gly Ser Pro  
 195 200 205  
 Ser Pro Ser Leu Leu Ser Ser Asp Leu Asn Leu Ser Tyr Thr Gly Asn  
 210 215 220  
 Asn Lys Leu His Ser Pro Ala Met Phe Leu Ser Ser Phe Asn Pro Arg  
 225 230 235 240  
 His His His Tyr Gln Ala Arg Ser Glu Asn Ser Asn Ile Ser  
 245 250 255  
 Cys Ser Leu Thr Met Gly Asn Pro Ala Met Val Gln Asp Lys Lys Lys  
 260 265 270  
 Ser Val Gly Ser Val Lys Thr His Gln Phe Val Leu Phe Gly Gln Pro  
 275 280 285  
 Ile Leu Thr Glu Gln Gln Val Met Asn Arg Lys Arg Phe Leu Glu Glu  
 290 295 300  
 Glu Ala Glu Ala Glu Glu Glu Lys Gly Leu Val Ala Arg Gly Leu Thr  
 305 310 315 320  
 Trp Asn Tyr Ser Leu Gln Gly Leu Glu Thr Gly His Cys Lys Val Phe  
 325 330 335  
 Met Glu Ser Glu Asp Val Gly Arg Thr Leu Asp Leu Ser Val Ile Gly  
 340 345 350  
 Ser Tyr Gln Glu Leu Tyr Arg Lys Leu Ala Glu Met Phe His Ile Glu  
 355 360 365  
 Glu Arg Ser Asp Leu Thr His Val Tyr Arg Asp Ala Asn Gly  
 370 375 380  
 Val Ile Lys Arg Ile Gly Asp Glu Pro Phe Ser Asp Phe Met Lys Ala  
 385 390 395 400  
 Thr Lys Arg Leu Thr Ile Lys Met Asp Ile Gly Gly Asp Asn Val Arg  
 405 410 415  
 Asn Ile Phe Ser Phe Leu Gln Ile Ile Phe Phe Phe Phe Phe Lys  
 420 425 430  
 Lys Leu Ile Phe Phe Leu Phe Ser Asp Gly Tyr Tyr Thr His  
 435 440 445  
 Ile His Ile Tyr Ser Ser Asp Tyr Ile Tyr Ile Tyr Ile Tyr Tyr  
 450 455 460  
 Ile Ser Asn Ile Lys Lys Lys Leu Ile Leu Phe Ile Leu Leu Ser Leu  
 465 470 475 480  
 Lys Thr Val Asn Ser Phe Ile Ser Ser Met Met Lys Lys Leu Ile Gln  
 485 490 495  
 Leu Ser Phe Thr Val Met Ile Ile Phe Thr Ile Leu Val Leu Gly Val  
 500 505 510  
 Val Ala Asn Glu Gly Leu Gly Lys Pro Lys Lys Gln Cys Asn Glu Ile  
 515 520 525  
 Leu Lys Gln Ser Asn Cys Val Ala Ala Glu Cys Asp Ser Met Cys Val  
 530 535 540  
 Lys Lys Arg Gly Lys Gly Ala Gly Tyr Cys Ser Pro Ser Lys Lys Cys  
 545 550 555 560  
 Tyr Cys Tyr Tyr His Cys Pro  
 565

&lt;210&gt; 508

&lt;211&gt; 776

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..776:Ceres Seq. ID 1941142

&lt;400&gt;508

agtcgctcc ttcctgcgc gcgcgcgcgc gcgcgcgcgc cgtacacac gctcgcgcgc 60  
 gctcgcgcgc tccgcgttc cgcaggaag ttggtgaag ttcttgagg ctggaacat 120  
 ggtcgtgaag cgcgaactt cgcgttcag cgcaggaag attatctg ggaagcat 180  
 tagattatc cgtcgtgatt ctgagttt ccttttgc aactcgaat gcaagcgtc 240  
 tttccacac cgtcgtgaag ctgcaagat taccctgaca gcaatgaca ggaagcaga 300  
 caagaagat attcatgtg aagcgtaaa gaagaggcg cgcgcaccca agaagccata 360  
 ctccagta attggtgtg ctctcttga agtaatccag aagaagag ctgagaagc 420  
 agagctccg gatgctgcta gaagagcgc tctcttgag atcaaggag cgtacaagaa 480  
 gaccaagat gagaagaag cgaagagcg gagggtgag aagtcocaga agacgcagc 540  
 aaggggtcg gtccagaag gtcccaagg cccaagtg gcgggcgtg gtgggaagcg 600  
 ctgaagaac ttagtgtct tctcgacat tgcagtcgt ccttagocaa agccacttc 660  
 gtgaacccg tgtgaattt gcaagacta ttcaagcgt gcttgctg gtaataacc 720  
 atggcaagag aacggattt tattatgc tgaataaat gaccgttcatt attctt 776

&lt;210&gt; 509

&lt;211&gt; 161

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..161:Ceres Seq. ID 1941143

&lt;400&gt;509

Met Val Leu Lys Thr Glu Leu Cys Arg Phe Ser Gly Gln Lys Ile Tyr 1  
 1 5 10 15  
 Pro Gly Lys Gly Ile Arg Phe Ile Arg Ala Asp Ser Gln Val Phe Leu 20  
 20 25 30  
 Phe Ala Asn Ser Lys Cys Lys Arg Tyr Phe His Asn Arg Leu Lys Pro 35  
 35 40 45  
 Ala Lys Leu Thr Trp Thr Ala Met Tyr Arg Lys Gln His Lys Lys Asp 50  
 50 55 60  
 Ile His Ala Glu Ala Val Lys Lys Arg Arg Arg Ala Thr Lys Lys Pro 65  
 65 70 75 80  
 Tyr Ser Arg Ser Ile Val Gly Ala Ser Leu Glu Val Ile Gln Lys Lys 85  
 85 90 95  
 Arg Ala Glu Lys Pro Glu Val Arg Asp Ala Ala Arg Glu Ala Ala Leu 100  
 100 105 110  
 Arg Glu Ile Lys Glu Arg Ile Lys Lys Thr Lys Asp Glu Lys Lys Ala 115  
 115 120 125  
 Lys Lys Ala Glu Val Ser Lys Ser Gln Lys Thr Gln Thr Lys Gly Ala 130  
 130 135 140  
 Val Gln Lys Gly Ser Lys Gly Pro Lys Leu Gly Gly Gly Gly Lys 145  
 145 150 155 160  
 Arg

&lt;210&gt; 510

&lt;211&gt; 106

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..106:Ceres Seq. ID 1941144

&lt;400&gt;510

Met Tyr Arg Lys Lys Gln His Lys Lys Asp Ile His Ala Glu Ala Val Lys 1  
 1 5 10 15  
 Lys Arg Arg Arg Ala Thr Lys Lys Lys Pro Tyr Ser Arg Ser Ile Val Gly 20  
 20 25 30  
 Ala Ser Leu Glu Val Ile Gln Lys Lys Arg Ala Glu Lys Pro Glu Val 35  
 35 40 45  
 Arg Asp Ala Ala Arg Glu Ala Ala Leu Arg Glu Ile Lys Glu Arg Ile 50  
 50 55 60  
 Lys Lys Thr Lys Asp Glu Lys Lys Ala Lys Lys Ala Glu Val Ser Lys 65  
 65 70 75  
 Ser Gln Lys Thr Gln Thr Lys Gly Ala Val Gln Lys Gly Ser Lys Gly 80  
 80 85 90 95  
 Pro Lys Leu Gly Gly Gly Lys Arg 100  
 100 105

&lt;210&gt; 511

&lt;211&gt; 1024

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..1024:Ceres Seq. ID 1942975

&lt;400&gt;511

ataatctc cgtctccct tgcgcgttc ctgtcgcgc gcgcgtccc tcccaacca 60  
 ctgctcaaa accttagagc tctcgtcgc tgcctcgc gcgcggagcg agtcccat 120  
 ctctccat caaccatgc gaccagatc agcaagaaga agaagttcgt cagcgcagc 180  
 gtittacg cagagtcac tagatgctg acgcggagc tggcgagga cggctactt 240  
 ggcgtgagg tgcggtcac gccgatgagc aagagatca tcatcgcgc cagcgcagc 300  
 cagacgcgc tgcggagaa gggcgttagg atcaggagc tcatctcgt cgtccagag 360  
 agttcaat tccctgaaa tgcgttagg cttatgccg agaagtcgt caacctggg 420  
 ctgtgcca tgcgcagcg cagtccttc cgttacaagc tctcgttgg cttgcgtc 480  
 cgcagggtt gctacgtgt tctctgtat gttatgaga gtgtgcaa ggtgtgag 540  
 gtatcgtga gtggaagct cagggccaa agacaaagt ccatgaagt caagatggc 600  
 tacatgct catctgtca accagtgaac ggtacattg actcgttgt gagacagt 660  
 ctctcagc aggggttct tggatcaag gtgaagatca cctctgact ggaccagag 720  
 ggcagggtt gccgatcac tctcttcgc gactcgtga cctccacac cccgaagac 780  
 gagcagagc cgcgcctcc ggtctggca cctctgagc tgaaggtcc caactgttc 840  
 ctgaaatgg cagttgatc ggtctcaga taggttcta tctacgata aggcacatc 900  
 taaggcagt ggtacctta atcttatgg acagatcta tgttgttct ctgtggcact 960  
 ttgtttaag ctaagtttg cggagacaag atactatat atgttccga tgaattatg 1024  
 gtic

&lt;210&gt; 512

&lt;211&gt; 274

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..274:Ceres Seq. ID 1942976

&lt;400&gt;512

Ile Asn Ser Ser Ala Pro Leu Arg Arg Phe Leu Ser Arg Arg Arg Ser

1 Pro Pro Asn Pro Leu Ala Gln Asn Pro Arg Ala Leu Ala Cys Ser Pro 15  
 20 20 25 30  
 Pro Pro Glu Arg Ala Pro His Leu His Pro Ser Thr Met Ala Thr 45  
 35 40 45  
 Gln Ile Ser Lys Lys Lys Phe Val Ser Asp Gly Val Phe Tyr Ala 60  
 50 55 60  
 Glu Leu Asn Glu Met Leu Thr Arg Glu Leu Ala Glu Asp Gly Tyr Ser 80  
 65 70 75 80  
 Gly Val Glu Val Arg Val Thr Pro Met Arg Thr Glu Ile Ile Ile Arg 95  
 85 90 95  
 Ala Thr Arg Thr Gln Asn Val Leu Gly Glu Lys Gly Arg Arg Ile Arg 110  
 100 105 110  
 Glu Leu Thr Ser Val Val Gln Lys Arg Phe Asn Phe Pro Glu Asn Gly 125  
 115 120 125  
 Val Glu Leu Tyr Ala Glu Lys Val Val Asn Arg Gly Leu Cys Ala Ile 140  
 130 135 140  
 Ala Gln Ala Glu Ser Leu Arg Tyr Lys Leu Leu Gly Gly Leu Ala Val 160  
 145 150 155 160  
 Arg. Arg Ala Cys Tyr Gly Val Leu Arg Tyr Val Met Glu Ser Gly Ala 175  
 165 170 175  
 Lys Gly Cys Glu Val Ile Val Ser Gly Lys Leu Arg Ala Gln Arg Ala 190  
 180 185 190  
 Lys Ser Met Lys Phe Lys Asp Gly Tyr Met Ile Ser Ser Gly Gln Pro 205  
 195 200 205  
 Val Asn Glu Tyr Ile Asp Ser Ala Val Arg His Val Leu Leu Arg Gln 220  
 210 215 220  
 Gly Val Leu Gly Ile Lys Val Lys Ile Met Leu Asp Trp Asp Pro Lys 240  
 225 230 235 240  
 Gly Lys Val Gly Pro Ile Thr Pro Leu Pro Asp Leu Val Thr Ile His 255  
 245 250 255  
 Thr Pro Lys Asp Glu Asp Glu Pro Arg Pro Pro Val Leu Ala Pro Pro 270  
 260 265 270  
 Glu Val

<210> 513  
 <211> 229  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..229:Ceres Seq. ID 1942977

<400>513  
 Met Ala Thr Gln Ile Ser Lys Lys Lys Phe Val Ser Asp Gly Val 15  
 5 10 15  
 Phe Tyr Ala Glu Leu Asn Glu Met Leu Thr Arg Glu Leu Ala Glu Asp 30  
 20 25 30  
 Gly Tyr Ser Gly Val Glu Val Arg Val Thr Pro Met Arg Thr Glu Ile 45  
 35 40 45  
 Ile Ile Arg Ala Thr Arg Thr Gln Asn Val Leu Gly Glu Lys Gly Arg 60  
 50 55 60  
 Arg Ile Arg Glu Leu Thr Ser Val Val Gln Lys Arg Phe Asn Phe Pro 80  
 65 70 75 80  
 Glu Asn Gly Val Glu Leu Tyr Ala Glu Lys Val Val Asn Arg Gly Leu 95  
 85 90 95  
 Cys Ala Ile Ala Gln Ala Glu Ser Leu Arg Tyr Lys Leu Leu Gly Gly 110  
 100 105 110

Leu Ala Val Arg Arg Ala Cys Tyr Gly Val Leu Arg Tyr Val Met Glu 120  
 115 120 125  
 Ser Gly Ala Lys Gly Cys Glu Val Ile Val Ser Gly Lys Leu Arg Ala 140  
 130 135 140  
 Gln Arg Ala Lys Ser Met Lys Phe Lys Asp Gly Tyr Met Ile Ser Ser 155  
 145 150 155 160  
 Gly Gln Pro Val Asn Glu Tyr Ile Asp Ser Ala Val Arg His Val Leu 175  
 165 170 175  
 Leu Arg Gln Gly Val Leu Gly Ile Lys Val Lys Ile Met Leu Asp Trp 190  
 180 185 190  
 Asp Pro Lys Gly Lys Val Gly Pro Ile Thr Pro Leu Pro Asp Leu Val 205  
 195 200 205  
 Thr Ile His Thr Pro Lys Asp Glu Asp Glu Pro Arg Pro Pro Val Leu 220  
 210 215 220  
 Ala Pro Pro Glu Val 225

<210> 514  
 <211> 206  
 <212> PRT  
 <213> Arabidopsis thaliana

<220>  
 <223> any n or Xaa = unknown

<223> LOCATION: 1..206:Ceres Seq. ID 1942978

<400>514  
 Met Leu Thr Arg Glu Leu Ala Glu Asp Gly Tyr Ser Gly Val Glu Val 15  
 5 10 15  
 Arg Val Thr Pro Met Arg Thr Glu Ile Ile Ile Arg Ala Thr Arg Thr 30  
 20 25 30  
 Gln Asn Val Leu Gly Glu Lys Gly Arg Arg Ile Arg Glu Leu Thr Ser 45  
 35 40 45  
 Val Val Gln Lys Arg Phe Asn Phe Pro Glu Asn Gly Val Glu Leu Tyr 60  
 50 55 60  
 Ala Glu Lys Val Val Asn Arg Gly Leu Cys Ala Ile Ala Gln Ala Glu 80  
 65 70 75 80  
 Ser Leu Arg Tyr Lys Leu Leu Gly Gly Leu Ala Val Arg Arg Ala Cys 95  
 85 90 95  
 Tyr Gly Val Leu Arg Tyr Val Met Glu Ser Gly Ala Lys Gly Cys Glu 110  
 100 105 110  
 Val Ile Val Ser Gly Lys Leu Arg Ala Gln Arg Ala Lys Ser Met Lys 125  
 115 120 125  
 Phe Lys Asp Gly Tyr Met Ile Ser Ser Gly Gln Pro Val Asn Glu Tyr 140  
 130 135 140  
 Ile Asp Ser Ala Val Arg His Val Leu Leu Arg Gln Gly Val Leu Gly 160  
 145 150 155 160  
 Ile Lys Val Lys Ile Met Leu Asp Trp Asp Pro Lys Gly Lys Val Gly 175  
 165 170 175  
 Pro Ile Thr Pro Leu Pro Asp Leu Val Thr Ile His Thr Pro Lys Asp 190  
 180 185 190  
 Glu Asp Glu Pro Arg Pro Pro Val Leu Ala Pro Pro Glu Val 205  
 195 200 205

<210> 515  
 <211> 1069  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>

gagagctgct gaatcacaag gtgatgtgga taatgtctaa tatctacaa gtatgtctca 480  
 gtgtaact atcagttgag tttttttt tgggttgtt tttgtkata atgacttctt 540  
 ctgcaagat ggtgttgat tagttttt ttgcaata atcgtaataa gtttcgaaa 600  
 ctggagagt tgaagtgt gaacataga ttgtgttat cgaataaaa gttattctt 660  
 atgcctgca tgcattgtt gtgaattga ttttaatgc gtttttcag c 711

&lt;210&gt; 577

&lt;211&gt; 59

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..59:Ceres Seq. ID 1015866

&lt;400&gt;577

Met Gly Ser Arg Gly Ile Ile Asn Asp Lys Trp Ser Met Arg Ile Leu 15  
 1 Trp Gly Cys Ala Ile Gly Ser Ala Ile Gly Leu Tyr Met Val Ala Val 30

Glu Arg Gln Thr Gln Asn Arg Ala Arg Ala Met Ala Glu Ser Leu Arg 35  
 45 Ala Ala Glu Ser Gln Gly Asp Gly Asp Asn Val 50

&lt;210&gt; 578

&lt;211&gt; 47

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..47:Ceres Seq. ID 1015867

&lt;400&gt;578

Met Arg Ile Leu Trp Gly Cys Ala Ile Gly Ser Ala Ile Gly Leu Tyr 15  
 1 Met Val Ala Val Glu Arg Gln Thr Gln Asn Arg Ala Arg Ala Met Ala 30

Glu Ser Leu Arg Ala Ala Glu Ser Gln Gly Asp Gly Asp Asn Val 35  
 45

&lt;210&gt; 579

&lt;211&gt; 38

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..38:Ceres Seq. ID 1015868

&lt;400&gt;579

Ser Ser Gly His Arg Leu Ile Leu Val Ser Ser His Ser Leu Cys Ser 15  
 1 Phe Ile His Ser Ser Leu Ile Asn Ala Thr Asn Phe Val Ser Pro Leu 30

Ile Ser Ser Phe Leu Leu 35

<210> 580  
 <211> 432  
 <212> DNA  
 <213> Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..432:Ceres Seq. ID 1021371

&lt;400&gt;580

gtcatttct cgattctcac tcttcggtc acgactaat tctgagaaaa aaatgggat 60  
 cggaggat tatcaagat agtgggtcaa tgagattct atgggttgt gctatggaa 120  
 tggctatgg ttatcatg gtgtctgag agagacaac tcagaaacag gctcgtgcta 180  
 tggctgagag ttgagagct gctgaatcac aagtgatgg tgataatgc taatactac 240  
 caagtagtc tgaagtgaat acttcagtt gagttttt ttttggtgt tgttttgtt 300  
 aaatgactt cttctgcaa gatgtgtg atgatttc ttttgcga ataactgtaa 360  
 taagtttgc aaactggag agtgaagtt gctgaacata cgatttgtt tatcgaaaa 420  
 aaagtattt cc 432

&lt;210&gt; 581

&lt;211&gt; 76

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..76:Ceres Seq. ID 1021372

&lt;400&gt;581

His Phe Leu Asp Ser His Ser Ser Gly His Arg Leu Ile Leu Arg Lys 15  
 1 Lys Met Gly Ser Arg Gly Ile Ile Asn Asp Lys Trp Ser Met Arg Ile 30

Leu Trp Gly Cys Ala Ile Gly Ser Ala Ile Gly Leu Tyr Met Val Ala 35  
 40 Val Glu Arg Gln Thr Gln Asn Arg Ala Arg Ala Met Ala Glu Ser Leu 45

Arg Ala Ala Glu Ser Gln Gly Asp Gly Asp Asn Val 60  
 65 70

&lt;210&gt; 582

&lt;211&gt; 59

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..59:Ceres Seq. ID 1021373

&lt;400&gt;582

Met Gly Ser Arg Gly Ile Ile Asn Asp Lys Trp Ser Met Arg Ile Leu 15  
 1 Trp Gly Cys Ala Ile Gly Ser Ala Ile Gly Leu Tyr Met Val Ala Val 30

Glu Arg Gln Thr Gln Asn Arg Ala Arg Ala Met Ala Glu Ser Leu Arg 35  
 40 Ala Ala Glu Ser Gln Gly Asp Gly Asp Asn Val 45

WO 00/40695

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..43:Ceres Seq. ID 1022580

&lt;400&gt;586

Met Val Ser Tyr Met Ser Thr Leu Arg Ser Leu Tyr Tyr Arg Glu Ser  
1 5 10 15  
Cys Cys Arg Arg Ser Gln Val Ser Gly Val Gly Phe Val Lys Ile Leu  
20 25 30  
Leu Thr Leu Phe Glu Phe His Val Tyr Cys Gly  
35 40

&lt;210&gt; 587

&lt;211&gt; 40

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..40:Ceres Seq. ID 1022581

&lt;400&gt;587

Met Val Arg Leu Tyr Lys Lys Pro Arg Asn Arg Ile Arg Gly  
1 5 10 15  
Lys Ala Lys Lys Lys Ser Lys Thr Ser Ser Gly Leu Ser Ser Lys  
20 25 30  
Ser Ala Lys Lys Asn Pro Trp Val  
35 40

&lt;210&gt; 588

&lt;211&gt; 643

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..643:Ceres Seq. ID 1024240

&lt;400&gt;588

agactcagtg aggtattttt aggcgatttt ctgcgaaaaa tgcgttgagg aaatggaga  
cttctatgag gtataccagc aattccaagt ctatgaagt tcatgcaaaa gagaaggttc  
cgggaactc aaaaacccat ttacagctc atgagagtt agatactgga actggggttc  
cgagttact ctgtgcgagtg attagacact ttttctcga ggccttgagg taggatgca  
ttatgataag cgcaaaagc ttcggtgtct tgcacgcgga aaaaagagt ttcctgtaag  
agctgataag cgttaacct ttaatttaa aggcggtgt gatattgac agacttaaa  
tcagaagaac cccaaggag cagcagaatt tgcctggaac ataaggtatt tcaaggaga  
tcagatgta cggatcaaa tggtacga aatgtttagt aaggtccctt atatgcagat  
tagagaaaac atgttgact tcacgcgga catgaaggga aaatgaaact tgcggtatga  
ctgtcaactg catittttc aatcatcgc tgagaaagt attgatacca ctgctgatga  
acacattta attotaccaa ttaataaat tcagagatct tcc

&lt;210&gt; 589

&lt;211&gt; 59

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

PCT/US00/00466

WO 00/40695

50 55

&lt;210&gt; 583

&lt;211&gt; 47

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..47:Ceres Seq. ID 1021374

&lt;400&gt;583

Met Arg Ile Leu Trp Gly Cys Ala Ile Gly Ser Ala Ile Gly Leu Tyr  
1 5 10 15  
Met Val Ala Val Glu Arg Gln Thr Gln Asn Arg Ala Arg Ala Met Ala  
20 25 30  
Glu Ser Leu Arg Ala Ala Glu Ser Gln Gly Asp Gly Asp Asn Val  
35 40

&lt;210&gt; 584

&lt;211&gt; 286

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..286:Ceres Seq. ID 1022578

&lt;400&gt;584

caccattgga tggttcggt atacaagctg aaactccca ggaacaggat tctgtgtaaa  
gcaagaac tgaactcga acaagctct ggttgagtt caaatcagc gaagaagac  
ccgtgggttt agaaatccca gaatttcaa tgttgatgg tgaattatat gtaactctt  
aggctcttt attatcaga aagttgtgc aagcagcagc aagtttctgg tgraggattc  
gtaaaaattt tgttaacttt attcgaattt catgtttatt ggggag

&lt;210&gt; 585

&lt;211&gt; 43

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..43:Ceres Seq. ID 1022579

&lt;400&gt;585

His His Trp Met Val Arg Leu Tyr Lys Leu Lys Pro Pro Arg Asn Arg  
1 5 10 15  
Ile Arg Gly Lys Ala Lys Lys Leu Lys Ser Lys Thr Ser Ser Gly Leu  
20 25 30  
Ser Ser Lys Ser Ala Lys Lys Asn Pro Trp Val  
35 40

&lt;210&gt; 586

&lt;211&gt; 43

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

301

SUBSTITUTE SHEET (RULE 26)

302

SUBSTITUTE SHEET (RULE 26)

&lt;223&gt; LOCATION: 1..59:Ceres Seq. ID 1024241

<400>589  
Met Glu Thr Ser Met Arg Tyr Thr Ser Asn Ser Lys Ser Met Lys Ile  
1 15  
His Ala Lys Glu Lys Val Pro Val Asn Ser Lys Thr His Leu Gln Leu  
20 30  
His Gly Glu Leu Asp Thr Gly Thr Gly Ala Pro Ser Tyr Phe Cys Ala  
35 40 45  
Met Ile Arg His Phe Phe Pro Glu Ala Leu Gly  
50 55

&lt;210&gt; 590

&lt;211&gt; 55

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..55:Ceres Seq. ID 1024242

<400>590  
Met Arg Tyr Thr Ser Asn Ser Lys Ser Met Lys Ile His Ala Lys Glu  
1 10 15  
Lys Val Pro Val Asn Ser Lys Thr His Leu Gln Leu His Gly Glu Leu  
20 25 30  
Asp Thr Gly Thr Gly Ala Pro Ser Tyr Phe Cys Ala Met Ile Arg His  
35 40 45  
Phe Phe Pro Glu Ala Leu Gly  
50 55

&lt;210&gt; 591

&lt;211&gt; 47

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..47:Ceres Seq. ID 1024243

<400>591  
Met Asp Phe Lys Glu Asp Gln Asp Val Arg Ile Lys Val Gly Tyr Glu  
1 5 10 15  
Met Phe Asp Lys Val Pro Tyr Met Gln Ile Arg Glu Asn Asn Trp Thr  
20 25 30  
Leu Asn Ala Asn Met Lys Gly Lys Trp Asn Leu Arg Tyr Asp Leu  
35 40 45

&lt;210&gt; 592

&lt;211&gt; 729

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..729:Ceres Seq. ID 1026562

&lt;400&gt;592

gttaacta taattgac cctaaatct atctcatct ccatcaatc ttctcttt  
aaatccaa atcttaagt tactttcta tctacattt caattcaat ggtttgatt  
ataactgt cgtctacc cagatccga gcatctccg gatccgac ccttaaccg  
gatcaaac gtaagaatc tgcgtctgg tgggtctctc ttctcggtt accctcgat  
ccgattacc tcaacatga asgtctatg tccacgtga atccgataa aacgatatt  
tcoggtcgg gtaaaagt ttctcggtt tctttacg aggegaagc taagcgttg  
agaggaana ccgcgaagc ttcaactc catgagtaa tgtatcact cgcattgct  
tctcaactt cgtctgat cccggcga gtcaggtt gaacgatt cgggtcaagt  
ctgagaga atcaatkaat gaataaga ttaatttt ttcttaast taagtgtct  
tgaagatgc agacatctt ctatcggtc tcttttcca ttgaacgt tgggtggt  
tggctcttg tgggtgaa actctggtt cctttctgt tctctgtaa tcaacgttg  
gattccact tgtattttt agacatgtt tccctagtta gctctttat cttatctaa  
attaaggc

&lt;210&gt; 593

&lt;211&gt; 153

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..153:Ceres Seq. ID 1026563

<400>593  
Val Asn His Ile Asn Leu Thr Leu Lys Ser Ile Ser Phe Ser His Gln  
1 5 10 15  
Ser Ser Ser Phe Lys Ser Leu Asn Leu His Val Thr Phe Leu Ser Thr  
20 25 30  
Phe Ser Leu Ser Met Ala Leu Ile Ile Thr Cys Ser Ala Leu Pro Thr  
35 40 45  
Ile Arg Ala Ser Ser Gly Ser Leu Asn Pro Asp Gln Asn Arg  
50 55 60  
Lys Lys Ser Ala Ala Trp Trp Ala Pro Leu Phe Gly Leu Pro Ser Asp  
65 70 75 80  
Pro Asp Tyr Leu Asn Ile Glu Ser Ser Cys Ser Thr Val Asn Pro Asp  
85 90 95  
Lys Thr Asp Ile Ser Gly Ser Gly Gln Lys Phe Arg Arg Gly Cys Phe  
100 105 110  
Thr Glu Glu Lys Ala Lys Gln Leu Arg Arg Lys Thr Ala Glu Ala Ser  
115 120 125  
Thr Phe His Asp Val Met Tyr His Ser Ala Ile Ala Ser Arg Leu Ala  
130 135 140  
Ser Asp Ile Thr Gly Arg Val Glu Asp  
145 150

&lt;210&gt; 594

&lt;211&gt; 117

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..117:Ceres Seq. ID 1026564

<400>594  
Met Ala Leu Ile Ile Thr Cys Ser Ala Leu Pro Thr Ile Arg Ala Ser  
1 5 10 15  
Ser Gly Ser Gly Ser Leu Asn Pro Asp Gln Asn Arg Lys Lys Ser Ala  
20 25 30

Ala Trp Trp Ala Pro Leu Phe Gly Leu Pro Ser Asp Pro Asp Tyr Leu  
35 40 45  
Asn Ile Glu Ser Ser Cys Ser Thr Val Asn Pro Asp Lys Thr Asp Ile  
50 55 60  
Ser Gly Ser Gly Gln Lys Phe Arg Arg Gly Cys Phe Thr Glu Glu Lys  
65 70 75 80  
Ala Lys Gln Leu Arg Arg Lys Thr Ala Glu Ala Ser Thr Phe His Asp  
85 90 95  
Val Met Tyr His Ser Ala Ile Ala Ser Arg Leu Ala Ser Asp Ile Thr  
100 105  
Gly Arg Val Glu Asp  
115

&lt;210&gt; 595

&lt;211&gt; 484

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..484:Ceres Seq. ID 1026648

&lt;400&gt;595

agagatcacca aaattaggtt tttaattga tagagaagat gaattccaga agcttgagg  
120 agtctggcc ttctacact gaaacacatc gaacgaagg ttgcacttca  
180 tagttatcat cgcagacatc gtgtgtttga tatgttcgat ttgatcaac tgggtgtct  
240 tagctctggt gactctggtt gggatcagat tgcggtggtta tagccacttc ttgtggaag  
300 ggaattccc ggcaggtttt gggatccgc ttgtgtggtt tctctgcgat ctcaagatgt  
360 tttagttgat gctcacaga agcatgagga gagagatgaa gaacttggat agagggccat  
420 tgttgagctt ctttgaagc agcagcagat atctctatag aattgtctt gatcttctt  
480 attgcttctt ggaattggtt tttagactcc agtttgtaat tacttcatgg aatcgttgtt  
484 tgcg

&lt;210&gt; 596

&lt;211&gt; 112

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..112:Ceres Seq. ID 1026649

&lt;400&gt;596

Met Asn Phe Arg Ser Phe Glu Glu Phe Trp Pro Phe Tyr Met Met Gln  
1 5 10 15  
His Ser Asn Pro Ser Thr Arg Arg Leu His Phe Ile Gly Ile Ile Ala  
20 25 30  
Ser Ile Val Ala Leu Ile Cys Ser Ile Leu Ile Asn Trp Trp Phe Leu  
35 40 45  
Ala Leu Val Pro Leu Leu Gly Tyr Gly Phe Ala Trp Tyr Ser His Phe  
50 55 60  
Phe Val Glu Gly Asn Val Pro Ala Ser Phe Gly His Pro Leu Trp Ser  
65 70 75 80  
Phe Leu Cys Asp Leu Lys Met Phe Ser Leu Met Leu Thr Gly Ser Met  
85 90 95  
Glu Arg Glu Met Lys Arg Leu Gly Lys Arg Pro Leu Leu Gln Leu Ser  
100 105 110

&lt;210&gt; 597

<211> 99  
<212> PRT  
<213> Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..99:Ceres Seq. ID 1026650

&lt;400&gt;597

Met Met Gln His Ser Asn Pro Ser Thr Arg Arg Leu His Phe Ile Gly  
1 5 10 15  
Ile Ile Ala Ser Ile Val Ala Leu Ile Cys Ser Ile Leu Ile Asn Trp  
20 25 30  
Trp Phe Leu Ala Leu Val Pro Leu Leu Gly Tyr Gly Phe Ala Trp Tyr  
35 40 45  
Ser His Phe Phe Val Glu Gly Asn Val Pro Ala Ser Phe Gly His Pro  
50 55 60  
Leu Trp Ser Phe Leu Cys Asp Leu Lys Met Phe Ser Leu Met Leu Thr  
65 70 75 80  
Gly Ser Met Glu Arg Glu Met Lys Arg Leu Gly Lys Arg Pro Leu Leu  
85 90 95  
Gln Leu Ser

&lt;210&gt; 598

&lt;211&gt; 98

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..98:Ceres Seq. ID 1026651

&lt;400&gt;598

Met Gln His Ser Asn Pro Ser Thr Arg Arg Leu His Phe Ile Gly Ile  
1 5 10 15  
Ile Ala Ser Ile Val Ala Leu Ile Cys Ser Ile Leu Ile Asn Trp Trp  
20 25 30  
Phe Leu Ala Leu Val Pro Leu Leu Gly Tyr Gly Phe Ala Trp Tyr Ser  
35 40 45  
His Phe Phe Val Glu Gly Asn Val Pro Ala Ser Phe Gly His Pro Leu  
50 55 60  
Trp Ser Phe Leu Cys Asp Leu Lys Met Phe Ser Leu Met Leu Thr Gly  
65 70 75 80  
Ser Met Glu Arg Glu Met Lys Arg Leu Gly Lys Arg Pro Leu Leu Gln  
85 90 95  
Leu Ser

&lt;210&gt; 599

&lt;211&gt; 563

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;223&gt; any n or Xaa = unknown

&lt;223&gt; LOCATION: 1..563:Ceres Seq. ID 1027881